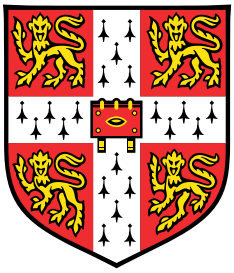
Vision-Based Over-Height Vehicle Detection for Warning Drivers



Bella Nguyen

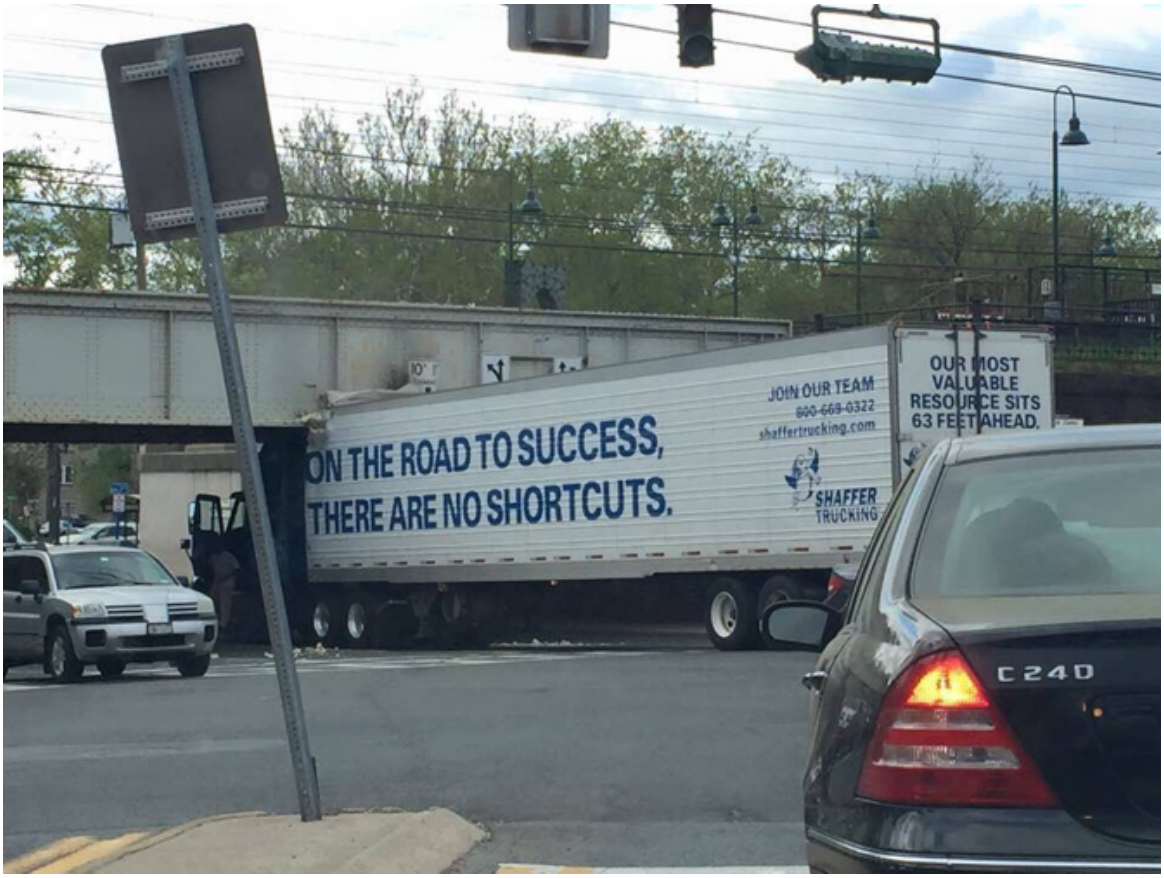
Laing O'Rourke Centre for Construction Engineering & Technology
Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Darwin College

November 2017

To those unfortunate lorry drivers . . .



Source: Unknown

Declaration

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university or institution for any degree, diploma, or other qualification.

In accordance with the Department of Engineering guidelines, this thesis does not exceed 65,000 words, and it contains less than 150 figures.

Signed: **Date:**

Bella Nguyen
Cambridge

Acknowledgements

Firstly, I would like to thank my PhD supervisor, Dr. Ioannis Brilakis, for his commitment and guidance throughout my studies. His 'open door' policy provided me with the support to excel and contributed to many of the successes during my research. All of this would not be possible without him. Special thanks are owed to Dr. Patricio A. Vela for taking a non-computer science student under his wings and having to teach the basics of computer vision to a civil engineer; all in layman's terms. We will have shaken beef again. Many thanks to Dr. Mohammed Elshafie for his support and reasoning throughout my PhD. I could not have done it without him.

Particular thanks are due to my lab mates, the Construction Information Technology (Cambridge) and the Intelligent Vision and Automation Group (Georgia Tech) for their advice and the many, many long coffee breaks spent at Starbucks.

Thank you to my family for their enthusiasm and cheer throughout, and especially to my sisters and mom for their continual love, care and support. Most importantly, this PhD would not be possible without the support of my fellow Darwinians for the many hours spent rowing, laughing and regaling. Thank you, Ying, Lizzy, Claire, Laurens, Florian, David, Philipp, Iva, Danny, Konstantin, Coto and Raj. In particular, thank you to Luke for his many hours spent proofreading my written work.

Finally, I would like to thank my sponsors Cambridge Overseas Trust, Marie Curie International Research Staff Exchange Scheme (IRSES), London Underground (TfL) – Redbridge Council, Cambridge Centre for Smart Infrastructure (CSIC) and Georgia Institute of Technology. Extreme gratitude is owed to Paul Fidler for his many hours assisting me on the bridge strikes deployment and Professor Campbell Middleton for his guidance and support throughout my PhD.

Journal publications from data presented in thesis

- [*Under Review*] Nguyen, B., Brilakis, I., & Vela, P. A. (2017). Real-Time Validation of Vision-Based Over-Height Vehicle Detection System. *Journal of Advance Engineering Informatics*, *xx*(x), xxxxxx.
- Nguyen, B., Brilakis, I., & Vela, P. A. (2017). Optimized parameters for over-height vehicle detection under variable weather conditions. *Journal of Computing for Civil Engineering, 31*(5), 04017023.
- Nguyen, B. & Brilakis, I. (2016). Understanding the Problem of Bridge and Tunnel Strikes Caused by Over-Height Vehicles. *Transportation Research Procedia, 14*. 3915-3924.

Conference publications from data presented in thesis

- Nguyen, B. & Brilakis, I. (2017). Over-Height Vehicle Detection: Minimising misclassifications due to wind. *2nd Lean Construction (LC3) conference.* Session: Intelligent Transportation Systems (JC3). July 4–12, 2017. Heraklion, Crete, Greece.
- Nguyen, B. & Brilakis, I. (2016). Understanding the problem of bridge and tunnel strikes caused by over-height vehicles. *6th Transport Research Arena.* April 18–21, 2016. Warsaw, Poland.
- Nguyen, B., Brilakis, I., & Vela, P. A. (2016). Vision-Based Over-Height Vehicle Detection. *Transportation Research Board 95th Annual Meeting* (No. 16-3550). January 8, 2016. Washington, DC, USA.

Research disseminations

- Nguyen, B. & Brilakis, I. (2017). Vision-Based Over-Height Vehicle Detection. *2nd Lean Construction (LC3) conference.* Session: Intelligent Transportation Systems (JC3). July 4–12, 2017. Heraklion, Crete, Greece. *(Oral Presentation)*
- Nguyen, B., & Brilakis, I. (2016). Understanding the problem of bridge and tunnel strikes caused by over-height vehicles. *6th Transport Research Arena Conference.* Session: Bridges (TWE4). April 18–21, 2016. Warsaw, Poland. *(Poster Presentation)*
- Nguyen, B., Brilakis, I., & Vela, P. A. (2016). Vision-Based Over-Height Vehicle Detection. *Transportation Research Board 95th Annual Meeting.* Session: Critical Transportation Infrastructure Protection and Resilience (16-3550). January 10–14, 2016. Washington, DC, USA. *(Poster Presentation)*
- Nguyen, B. (2015). Future Technology Forum Dragons' Den Competition. *New Civil Engineer (NCE).* Session: October 1, 2015. London, UK. *(Competition)*

Scholarships, grants and awards obtained during PhD candidature

- Winner of Construction Innovation Awards, 2017 (Crete)
- Outstanding Student Research Project, Fiatech (CETI AWARD), project: Autonomous Machine Vision Bridge and Tunnel Strike Prevention System (2016)
- Rex Moir Travel Grant (2015)
- Marie Curie International Research and Staff Exchange Scholarship (2014-15)
- Cambridge Overseas Trust Scholarship (2014-16)
- Cambridge Engineering Department Travel Grant (2014)
- Darwin College Travel Grant (2014)
- London Underground (TfL) Research Fund, Grant Number RG59570 (2013-14)
- Cambridge Centre for Smart Infrastructure and Technology (CSIC) Grant, Project Number NMZJ/089, Task 21 (2013-14)

Abstract

Many older bridges and tunnels were constructed using standards by now many decades out-of-date, at a time when trucks and other large vehicles were smaller. A bridge or tunnel strike is an incidence in which a vehicle, typically a lorry (truck) or double-decker bus, tries to pass under a bridge or tunnel that is lower than its height, subsequently colliding with the structure. These strikes lead to an increased cost of bridge repairs, clogged up roadways and increased potential for catastrophic events: hazardous spillage and/or total collapse. Today, Network Rail reports on average a strike every 4.5 hours.

There are a number of reasons why strikes occur, and why drivers of heavy goods vehicles sometimes fail to recognise the warning signs, consequently striking the bridge or tunnel. At first glance, it may seem like the problem is a fairly easy one to solve; however, no matter how well planned the road system, human error is an ever-present risk.

The research proposes to address the problem of bridge and tunnel strike prevention and management. The intent of the research is to develop an affordable, reliable and robust early warning over-height detection system bridge-owners can implement at locations with high strike occurrences. The research aims to test and validate a novel vision-based system using a single camera to accurately detect over-height vehicles using a set of optimised parameters. The system uses a camera installed at the offending height, which acts as an "over-height plane" formed by the averages of the maximum allowable heights across all lanes in a given traffic direction. Any vehicle exceeding this plane is analysed within a region of interest using a trigger-based approach for accurate detection and driver warning. If the vehicle is deemed to be over-height, a warning is issued to the driver. As a result, prolonging life expectancy of structures while decreasing the cost of repairs, maintenance and inspections.

Table of contents

List of figures

List of tables

List of abbreviations and acronyms

BRISK	Binary robust invariant scalable keypoints
DTW	Dynamic trip wires
EIGEN	Shi and Tomasi's method (1994)
EWDS	Early warning detection system
HGV	Heavy goods vehicle
KLT	Kanade-Lucas-Tomasi
OH	Over-height
OHV	Over-height vehicle(s)
OHVD	Over-height vehicle detection
OHVS	Over-height vehicle strike[s]
Precision	Also known as detection rate
ROI	Region of interest
SIFT	Scale-invariant feature transform
Specificity	Also known as false alarm rate
SURF	Speeded-up robust features

Chapter 1

Current State of Practice

1.1 Introduction

An *over-height vehicle strike* (OHVS) is an incident in which a vehicle, typically a lorry (truck) or double-decker bus, tries to pass under a bridge or tunnel that is lower than its height, subsequently colliding with the structure. According to the US Federal Highway Administration, the third most common cause of bridge failure is collision damage when a vehicle or vessel hits a bridge (FHWA, 2013). Accidental collisions between *over-height vehicles* (OHV) and bridge superstructures are a frequent phenomenon occurring throughout transportation networks worldwide (Fu et al., 2004, El-Tawil et al., 2005, Xu et al., 2012). These strikes lead to traffic delays, damage to bridge structures, bridge closures and injuries. In the worst-case scenario, derailments, immediate collapse of bridge structures, and fatalities may occur (Ghose, 2009, Washington State Department of Transportation, 2013).

Managing OHVS requires attention in three domains: prevention (discouraging strikes in the first place); detection (accurately recording strikes that do occur); and reporting (efficiently communicating OHVS details to the relevant authorities). The latter two aspects of OHVS management are effectively covered by existing systems. Much existing OHVS technology is targeted towards preventing OHVS from occurring in the first place versus mitigating the impact. Very few systems are designed to mitigate OHVS impact, as asset owners are interested in protecting the structure and limiting any risk of structural instability.

Current prevention systems can be categorised into passive, sacrificial, and active types. Practitioners favour quick, cost-effective, and accessible passive methods such as signage, bridge markings, and flashing beacons as an initial attempt to warn drivers. Such passive systems are readily available, easily installed, and minimise additional infrastructure installation. However, they prevent only ~10-20% of strikes, meaning that complimentary systems are necessary for higher prevention rates (Cawley, 2002). Where strikes have persisted, practitioners may incorporate sacrificial or active systems. Sacrificial systems (also known as rigid passive systems) are ideal for asset owners as post-installation maintenance is minimal (this issue is further discussed in section 1.3.1.2).

Active systems, also known as *Early Warning Detection Systems* (EWDS), detect and notify vehicle operators ahead of the presence of low structures. Existing systems consist of a transmitter and a receiver, placed directly across the lane(s) of traffic with an inductive loop to detect presence of a vehicle in advance of the warning sign (TRIGG Industries International, 2015). Asset owners in the US, Australia, China, Canada and Netherlands have deployed active systems using laser or infrared light warning systems at low clearance locations (Alberta Infrastructure & Transportation, 2008, Sina, 2012, New York State, 2015, LaserVision, 2015 & Dutch Ministry of Infrastructure, & Environmental Department of Waterways and Public Works, 2015). However, at non-critical low height locations, most asset owners have chosen not to use EWDS, based on unfavourable cost-benefit analyses. The reported installation costs range in the hundreds of thousands of dollars, therefore limiting the widespread adoption of EWSD (Sandidge, 2012, Dai et al., 2015, Singhal, 2015).

The biggest issues for asset owners considering OHVS system installation are affordability and reliability, with minimal compromise of accuracy and performance. While many OHVS systems are currently on the market, none cover the three aspects of OHVS management affordably. Existing systems that target some aspect(s) of the OHVS problem are discussed, paying attention to their respective benefits and limitations. Therefore, the chapter aims to: (1) Synthesize the current state of practice in OHVD.

1.2 Nature and scope of the problem

The problem of bridge and tunnel strikes is first recognised in the literature in the 1970s, after an increase in the number of vehicle collisions with low railway bridges (Martin & Mitchell, 2004). There are a number of reasons why these strikes occur, and why drivers of *heavy goods vehicles* (HGV) may fail to recognise the warning signs, consequently striking the structure (Martin & Mitchell, 2004, Byrne, 2009, Ghose, 2009, Agrawal, 2011): a) drivers unaware of vehicle height, b) poor route planning, and c) inadequate warning of low bridges.

Bridge damage due to collisions of OHVS is a major issue occurring throughout transportation networks worldwide (Fu, Burhouse, & Chang, 2004, El-Tawil, Severino, & Fonseca, 2005, Agrawal, 2011). In Mississippi, heavy logging truck traffic has posed constant danger to the truss bridges common in rural areas (Hanchey & Exley, 1990). Multiple strikes were recorded, and the average cost of repair of the bridges is $200,000 USD at the time. In the UK, there are almost 10,000 railway bridges crossing over roadways. Of these, 3,400 (34%) are considered 'at risk', due to their low height (below 16' 6"/ 5.03 m) (Horberry, Halliday & Gale, 2002). Network Rail (2007a) reports that a vehicle strike with a railway bridge occurs on average once every four and a half hours. In Beijing, China, roughly 20% of bridge damage is caused by OHV (Sina, 2007).

Representative examples of recent severe OHVS come from Canada and the United Kingdom. Figure 1 shows a driver having left the bucket of a dump truck raised on a major skyway bridge connecting two cities in Ontario. Damages were severe, and inspectors feared

Fig. 1 July 2014 - Dump truck hits scaffolding after leaving buck raised on the Burlington Skyway Bridge in Ontario, Canada (SHAH, 2014)

a total bridge collapse. In Figure 2, a footbridge completely collapsed over a motorway in the county of Kent (UK) after being hit by a lorry carrying a digger. The collapse is severe, causing hours of traffic backup. Better preventative measures must be in place to prevent such events from occurring.

OHVS thus have great impact, not just in terms of damage to the infrastructure, but also on public transportation systems i.e. road and rail networks causing traffic delays and congestion. In the US, the largest bridge owner is the State of Texas, with 51,000 bridges and overpasses. According to the Texas Department of Transportation, a strike can average $180,000 USD when repair is required and can take a bridge or overpass out of service for up to a year. The state further reports that repair costs are easy to quantify, but the cost to the public from inconvenience, detours, and congestion is not (Meyer, 2013).

1.2.1 Bridge strike statistics

Shanafelt & Horn (1980) found that bridge engineers cited OHVS as the leading cause of damage (81%) to pre-stressed concrete bridges. A later study by the same authors over a five-year period found that 95% of damage to steel bridges in the US is caused by OHVS (Shanafelt & Horn, 1984). Harik et al. (1990) analysed US bridge failures over a 38-year period (1951-1988). Of the 79 bridge failures in the study, 11 (14%) were due to truck collisions with a bridge. The London Department of Transportation (1990) created an incident database, which contains details of 4285 separate OHVS. Strikes can range from a slight scrape to a direct collision which causes total bridge collapse. The UK Department of Transport reported that the number of OHVS at railway bridges increased from 729 in 1990 to 1,870 in 2004. The majority of damages generated costly repairs to the bridge infrastructure and caused delays on the transportation network.

Horberry, Halliday, & Gale (2002) describes an increase in OHVS from less than 300 per year in the 1970s to over 1700 per year from 2000 onwards. Fu et al (2004) reported that OHVS in Maryland had increased by 81% between 1995 and 2000. Of the 1,496 bridges in Maryland susceptible to impact, over 300 (ca. 20%) have been hit by OHV at some point.

Fig. 2 August 2016 - Footbridge collapses on a motorway in Kent, UK after being hit by a lorry carrying a digger (BBC News, 2016)

The study surveyed 29 states about OHVS, with 18 designating them a significant problem, although few were able to provide statistics. A study of Network Rail (UK) (Martin & Mitchell, 2004) reported that over the ten years leading up to the study, the number of strikes at railway bridges over roads ('underline') doubled. The Network Rail (UK) database contains 12,829 incidents for the period spanning 1995-2003. Approximately 50% of these incidents involved HGVs. A slight majority of bridges (1719; 55%) were only struck once during this period; 436 bridges (14%) were struck twice; 216 (7%) bridges were struck on three occasions and 760 (24%) bridges were struck four or more times. Of these, 109 bridges have been struck 21 times or more (Martin & Mitchell, 2004).

The most recent study on the seriousness of the OHVS problem across the US is conducted by Agrawal (2011). Of the 44 Departments of Transportation and 2 local authorities that responded, the majority of the states across the country consider OHVS to be a major problem (see Figure 3). It is unclear in the report how those state departments measured the level of 'seriousness' of the OHVS problem as opposed to other problems; there is a lack of baseline against which to measure seriousness or frequency.

For instance, Nebraska perceive OHVS to be a major problem even though there have been only 20 instances whereas Missouri has had 1691 instances and do not perceive OHVS as a serious problem. This difference is so extreme that it calls into question the whole study. The study hinges on the semantics of a single word ('serious'), necessarily treated differently

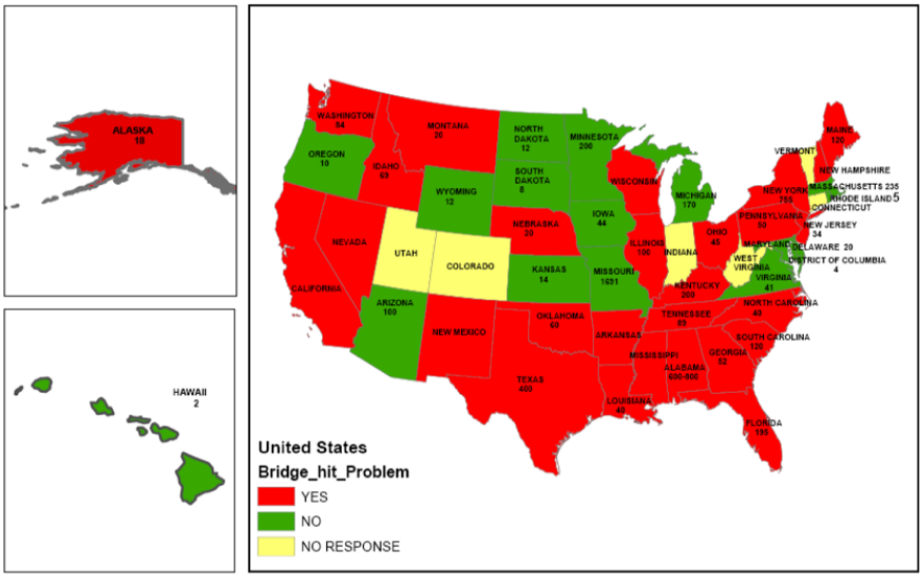


Fig. 3 Over-height vehicle strike problem in the United States, 2005-2008

by each respondent. The study is thus subjective, although it is of some use by gathering multiple statistics in one place. Based on OHVS data from 1980 - 2011, we are unable to find a unifying factor behind the increase in OHVS incidents. The significant point to take away is that strikes are still occurring, with the likely prospect of increasing frequency. The typical types of OHV on the roadways are shown in Figure 4.

The most prominent bridge strike champion, Network Rail, is an activist for bridge strike prevention and raising awareness. With over 32,000 bridges and tunnels within their network, bridges over roads were struck repeatedly in 2015/16, with 26 strikes at St Mildreds Road (A205) in Hither Green, London. Bridges in Thurlow Park Road in Tulse Hill and Lower Downs Road in Wimbledon followed by 22 and 16 bridge strikes respectively for the year ending March 2016. In 2014, the number of bridge strikes increased by 9.6% compared to the previous year (Network Rail, 2017). A reflection of the last ten years in Table 1 shows bridge strikes decreasing by 20% since 2006 but for underline bridge, the strikes are consistently ranging between 1500–1600 per year. For overline bridges, the average number of strikes remain at 115 with very few fluctuations year-on-year.

Bridge strike data from year 2015-16 compared to 2014-15 show a consistent number of strikes with underline and overline bridges of 1638 and 138 respectively. Despite efforts to raise awareness such as manuals for prevention and responding guides for drivers, the number of strikes remain between 1400 – 2100 over the last 10 years. The measureable performance loss shown in Table 2 due to strikes resulted in a 51.6% decrease in annual delays (minutes) from 2005 to 2011 however, the reported number of bridges increased by 22.0% in 2008 and decreased by 27.6% in 2011. The cost of delay shows that, although strikes are not decreasing, the annual cost of delays are decreasing by half (£4.8 mil) from 2005 to

Fig. 4 Typical over-height vehicles on the roadway

Year	Underline Bridge	Overline bridge
2006/07	2020	126
2007/08	2161	143
2008/09	1736	115
2009/10	1439	113
2010/11	1593	98
2011/12	1541	80
2012/13	1544	101
2013/14	1708	101
2014/15	1670	140
2015/16	1606	136

Table 1 Network Rail, ten years of bridge strike data from 2006–2016

	2005/06	2007/08	2010/11
Number of Reported Bridge Strikes	1930	2355	1704
Annual Delay (Minutes)	335,500 minutes	220,465 minutes	162,348 minutes
Average Delay per Bridge Strike	174 minutes	130 minutes	132 minutes
Annual Cost of Delay	£9.6m	£6.8m	£4.8m
Average Cost per Bridge Strike	£4,975	£3,900	£3,918

Table 2 Performance loss reported on Network Rail of reported bridge strikes ranging from 2005–2011.

2011. The assumptions drawn are strikes are better managed resulting in less time loss and less cost incurred by the owner.

1.2.2 Why do strikes occur?

As the previous section noted, incidences of OHV striking low bridges have increased considerably since the 1970s. Galer (1980) investigated two possible reasons for the accidents. These were drivers' knowledge of their vehicle heights, and drivers' understanding of the low bridge warning signs. Only 12% of drivers were correct in their estimate of their vehicle height and just 27% were within 3 in (76 mm) of the correct height.

In the UK, Chapter 4 of the Traffic Signs Manual states that "[t]he figures shown on the signs to indicate the available headroom should be at least 3 inches less than the measured height to allow a safety margin and should be expressed to the nearest multiple of 3 inches." In the US, some states post the actual vertical clearance on warning signs, while other states under-report the clearance by up to twelve inches (0.305 m). This can have negative effects, with drivers possibly ignoring signs in the knowledge that clearance is under-reported (Agrawal, 2011). The policy enforcement had a negative effect, resulting in drivers not believing the posted height of bridges or the height information (Martin & Mitchell, 2004).

1.2.3 Impacts & consequences

The traffic congestion and delays caused by OHVS fall at the lower end of the spectrum however, yet still disruptive. Strikes can bring traffic to a standstill for several hours while the OH truck is removed from the bridge or tunnel, and debris (if any) is cleared. At the low-to-medium end of the spectrum, the top of the vehicle may simply scrape the underside of the structure. With medium to severe incidents the structure itself may be damaged, i.e. breaking reinforcements, exposing pre-stressing steel and damaging concrete elements.

Bridges under railways are critical points on rail networks, and any congestion or blockage has ramifications for rail services. When OHVS occur, services are delayed until the bridge or tunnel has been inspected to determine if the structure has been compromised and whether it is safe to resume operations; traffic on and under the structure must be stopped pending the outcome of the structural inspection to check for instability (Network

Fig. 5 Over-height vehicle strike accidents by level of severity

Rail, 2007a). Speed restrictions or access prohibitions may be placed on the line while the inspections are completed. This causes revenue loss to railway companies, because track access agreements require reimbursement to train operating companies for any hindrance to track access.

Figure 5 schematises levels of severity of OHVS accidents on the transportation network (Nguyen et al., 2016). The spectrum ranges from minor (no casualties) to severe (many casualties). The severe end of the spectrum under 'fatalities' can involve an OHVS with a bridge or tunnel structure resulting in one or more casualties. Severity is defined in terms of human injury and damage to infrastructure. In most cases, railway services run above the bridge structure, posing a risk to road drivers themselves. OHVS have the potential to cause rail track displacement. In the worst-case scenario, this can cause a bridge to completely collapse. OHVS are thus a serious problem in the transportation and infrastructure industry, and adequate research directed at prevention has so far been lacking. At first glance, it may seem like the problem of OHVS is easily solved; however, no matter how well-planned a road system may be; driver error is an ever-present risk. Increasing fines and surveillance may be a partial solution to the OHVS problem, yet such strategies still do not eradicate the problem of human fallibility. This raises the question—*Why is the current state of practice insufficient for OHVS prevention?*

1.3 Current state of practice

In the UK, vehicle heights are typically under 4.5m, however, bridge structures are built to standards that are decades out of date and are often inadequate today causing an ongoing nuisance for asset owners. One of the earliest forms of the OHVS management system is designed over a century ago by the American engineer James H. Donaldson (1906).

The guard system shown in Figure 7 is intended to warn drivers that a train is about to pass into a tunnel or under a bridge. The guards consisted of a number of strips of flexible material attached to a wire stretched across the track striking the top of the train, and warning drivers to stop to allow for the train to pass. Over the years, this type of *over-height vehicle detection* (OHVD) and early warning system has evolved into the commonly used OHVS prevention tools still with us today. Many OHVS systems are available for consumer

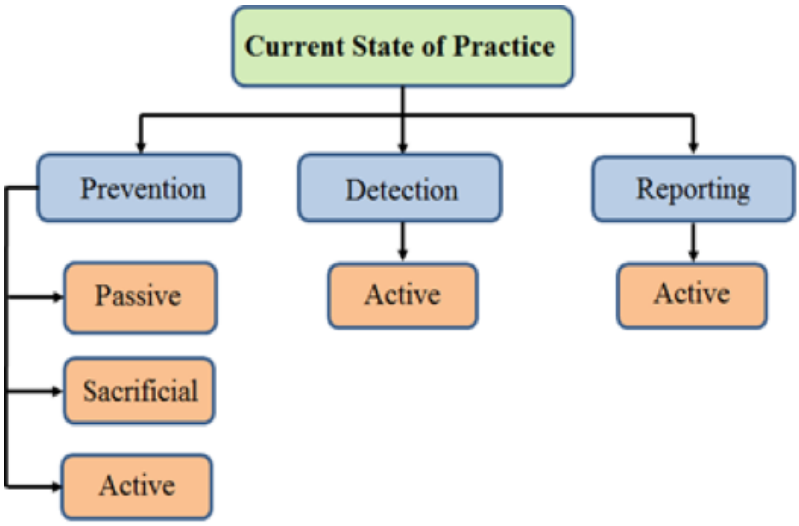


Fig. 6 Current state of practice for over-height vehicle strikes

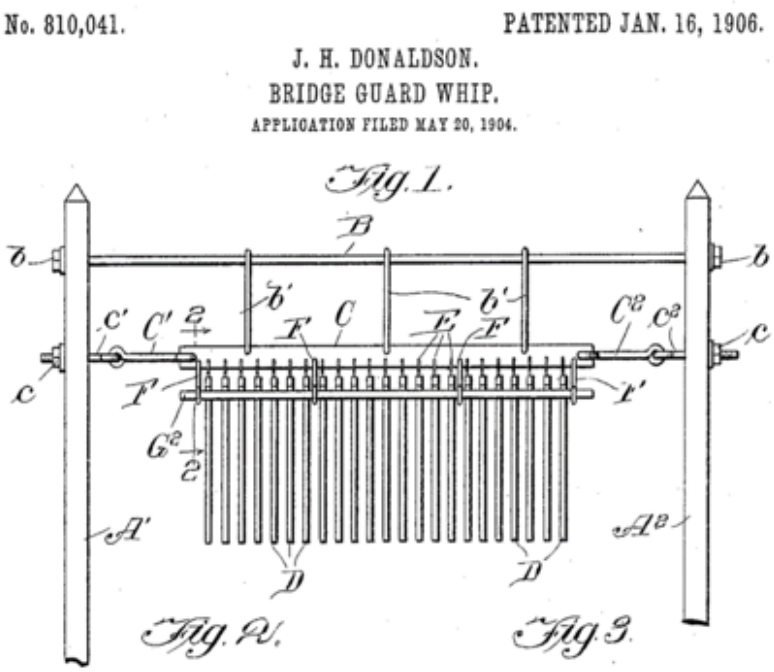


Fig. 7 1906 patent application, Bridge Guard Whip

purchase today. The current state of practice is best divided into three categories: prevention, detection and reporting (see Figure 6). To reiterate:

1. *Prevention* includes methods used to prevent OHVS from occurring;
2. *Detection* includes methods used to recognise a physical OHVS;
3. *Reporting* includes methods which report OHVS to authorities.

1.3.1 Prevention systems

Most existing OHVS technology on the market is targeted towards preventing OHVS from occurring. Within the prevention category, there are three basic OHVS protection subcategories: Passive (non-rigid), sacrificial (rigid) and active (rigid) systems. According to Cawley (2002: Pg. 9), passive signing is estimated to be 10-20% effective in preventing incidents; sacrificial systems are estimated to be 30-50% effective, and active warning systems are estimated to be 50-80% effective. The cost of installing an OH early warning detection system at bridges is typically much less than the cost of repairing damages from OHVS (Hanchey & Exley, 1990). In the remaining section, two types of prevention methods are described: non-rigid (passive) and rigid (sacrificial and active) methods.

1.3.2 Passive systems

Passive systems are the most common and cost-effective type of system that exist, due to their readily available supply and 'quick-fix' approach. These methods include static signage, variable message signs, flashing beacons and bridge markings. In the UK, the Department of Transport (2008; Ch. 3) requires all bridges under 16'6" / 5.03 m to be posted with a warning sign as shown in Figure 8. The red circles are considered 'prohibiting' signs while the red triangles are 'warning' signs.

According to the 2012 Transport Statistics of Great Britain, approximately 260,000 registered UK HGV drivers and 130,000 foreign HGV drivers enter the UK each year, with a total of 1.5 million journeys. Foreign drivers unfamiliar with prohibiting and warning signs may perceive both signs to be irrelevant, especially in instances where drivers need to make rapid decisions. Having two signs with different shapes on each bridge approaches only adds confusion and does not alleviate the problem of OHVS.

Variable message signs are another commonly used device to pre-empt drivers approaching low bridges (Fontaine, 2003, Martin & Mitchell, 2004, Network Rail, 2007a, Byrne, 2009, Agrawal, 2011, Thakuriah & Geers, 2013). A study by Horberry, Halliday & Gale (2002) tested different types of bridge markings.

The primary function of bridge markings is to make bridge openings appear smaller from a distance. The study used new designs of bridge markings and concluded that the new markings appeared more conspicuous, making drivers more reluctant to pass underneath.

(a)

(b)

Fig. 8 (a) Low bridge prohibiting and warning signs (UK); (b) Mandatory height limit for low bridge ahead, (middle) mandatory height limit ahead and indication of alternative route, (right) indicating a restriction on junction road ahead

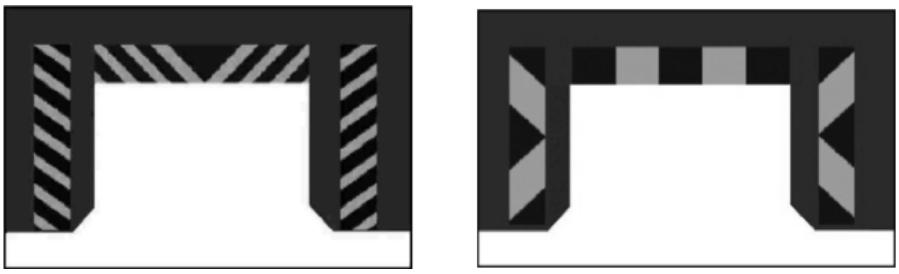


Fig. 9 (left) Old standard bridge marking design; (right) new bridge marking design

Fig. 10 Laser warning sign (pseudo holographic image that floats in mid-air, making it impossible for drivers to miss)

In 1993, as part of an experiment, the Department of Transportation repainted the markings on a bridge to camouflage the lower edge of the bridge deck (i.e. the top of the bridge) and the strike rate for that bridge tripled. The Department of Transportation concluded that bridges should be made to appear lower, and that this would reduce the accident rate (Figure 9). The method crucially relies on drivers to take appropriate precautions to make quick decisions leading to the preventative initiative addressing only part of the OHVS management problem. Additional preventative mechanisms are often required.

The latest innovation in passive warning is the softstop barrier system shown in Figure 10, developed by Laservision. The first system is installed at the Sydney Harbour Tunnel. The system creates the illusion of a solid surface that instantly blocks both lanes of traffic. The Softtop Barrier System produces a pseudo-holographic image that appears to float in mid-air, commanding the attention of the motorist making the 'STOP' message impossible to miss. The technology includes hydraulic water screens, critical pressure levels to mitigate distortion from wind currents, rapid start techniques, and monitor loops. The cost to purchase the rain curtains are in the range of $150,000-200,000 AUD with an annual maintenance cost from $25,000-35,000 AUD.

Other passive systems include flashing signs, flashing beacons (Fontaine, 2003; Hanchey & Exley, 1990, Mattingly, 2003) and flashing lights (Manjunathan, Albert, & Deeter, 2005) to notify over-height drivers that they should use an alternative route. These passive systems can be used in conjunction with other OHVS prevention systems. Flashing beacons are commonly used at low bridge approaches to warn drivers of an oncoming hazard, and are typically paired with other preventative methods such as the bridge markings described above.

Driver education, policies and manuals

At the policy level, asset owners have attempted to manage the problem of OHVS by implementing permits, axel load restrictions, fines, driver education and awareness programs, good practice manuals and newsletters. According to HM Revenue & Customs, vehicles taller than 3m in height and above must display a notice in the cab showing its full height. Best practices encourage drivers to check the height of the vehicle and display the height

Fig. 11 Transport Scotland raising awareness and its campaign to prevent bridge strikes (Transport Scotland, 2011)

in their cab before commencing a journey. The National Railway (UK) ran a campaign to increase awareness and offer advice regarding low bridges and preventing OHVS from 2002–2012. Extensive and comprehensive manuals exist: 'Good Practice Guides', protocols for passengers, professional drivers, transport managers and bridge owners.

Transport for Scotland pairs with the National Bridge Strike Prevention Group to develop a 'Strike it Out' Campaign (Donnelley, 2010). This group includes members of road and rail bridge organisations across the UK, as well as freight groups, police and policy makers with various transport bodies. The aim of the group is to raise OHVS awareness among those involved with driver training and management as shown in Figure 11. Although these strategies may not directly prevent OHVS from occurring, increased awareness plays a positive role and can be effective for passengers, professional drivers and transport managers.

Sacrificial systems

The second type of OHVS prevention scheme involves physical notification, i.e. sacrificial systems. Sacrificial systems consist of crash beams (also known as collision protection beams, impact beams, bridge bumpers, or cushion systems), hanging chains/strips/bells/headache bars, portal frames and road narrowing techniques such as speed bumps and rumble strips.

Crash beams are an effective method of mitigating structural damage to bridge structures from OHV impact (Qiao, Yang, & Mosallam, 2004, Sharma, Hurlebaus, & Gardoni, 2008; 2012). Crash beams act as a 'cushion' to the bridge structure (Yang & Qiao, 2010); energy transferred by the strike is dissipated by the beam therefore reducing damage to the main structure. An example of a crash beam is shown in Figure 12. This in turn protects the structure itself. However, risk of injuries and fatalities are still an exigent problem in such cases. Crash beams provide no advance warning to drivers, and can thus be viewed as a last resort for drivers who fail to notice the low bridge warnings. A personal interview with Mr. Ashok Parmar from London Underground reveals that a recent crash beam project cost £1,500,000 for each approach. Permit approvals can be a long and time-consuming process. Furthermore, crash beams require technical expertise from architects, engineers and construction managers (Sharma, Hurlebaus & Gardoni, 2008; 2012). Crash beams are

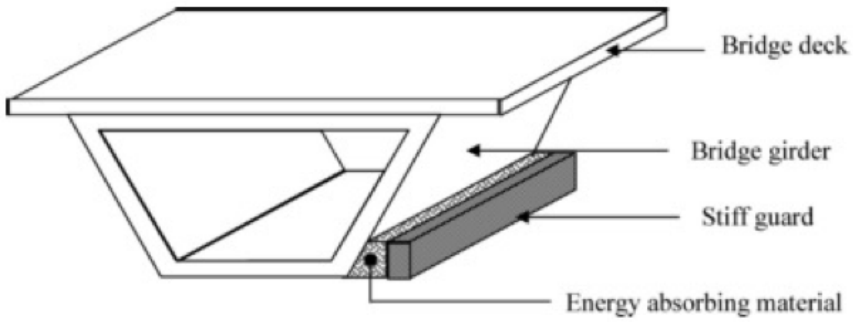


Fig. 12 A crash beam designed for a bridge bumper installation on bridge girder (Sharma et al., 2008)

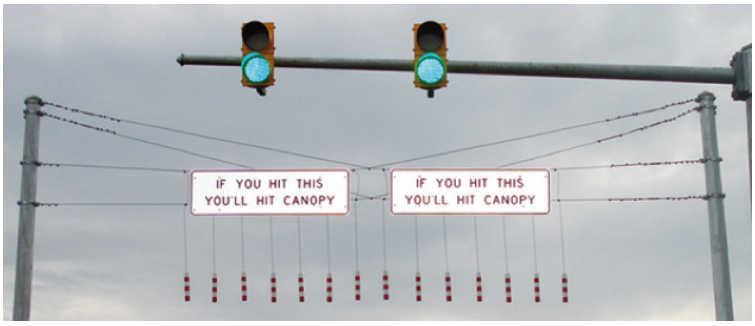
costly and an effective mitigation strategy but they too only solve part of the problem; the beams do not warn vehicle operators and are protective rather than preventative.

An alternative option is the use of road-narrowing (calming) techniques such as speed bumps, rumble strips and chicanes and metal hanging chains. These alternative options are commonly used as speed-reducing mechanisms. Similar to the 1906 invention, hanging chains are a commonly seen modern variant of the bridge guard whip. Reports have shown that chains are ineffective because it is hard for the drivers to hear the chains due to the loud engine (Sandidge, 2012). An example of a sacrificial system hanging chains is shown in Figure 13. Cawley (2002) estimates that sacrificial structures or metal strips cost $62,000 USD per installation. In Korea, hanging bells are used with a similar function (Byrne, 2009). However, weathering causes major damage to metal chains, and, in general, calming techniques require major road reconfiguration to be effective, making hanging chains an example of a non-ideal class of solutions to OHVS.

1.3.3 Active systems

The third OHVS prevention type is the intelligent transportation system or active warning alternative. At the basic end of active system types, Geographical Positioning Systems (GPS) are used by HGV drivers to indicate where low bridges are located. A small unit is installed in the vehicle cab, and as vehicles approach a low bridge, visual and audio warnings are activated within the cab if the vehicle is too high for the bridge (Martin & Mitchell, 2004). Transport for Scotland (Donnelley, 2010) has actively encouraged drivers not to trust satellite navigation, as the systems may not contain accurate bridge height information. Another active system is the EWDS.

The first OHVD system was patented by Lowry & Forster in 1981 (USA). The system uses sensors to detect the OHV, provides an audible warning, and guides the driver to an alternative route. The system has a pair of operating light sources and sensors spaced at a distance from each other in advance of the overhead structure (Figure 14).



(a)

(b)

Fig. 13 A sacrificial system that uses long hanging tubes and/or chains to create noise when a vehicle is in contact with the objects (FutureNet Solutions, 2017)

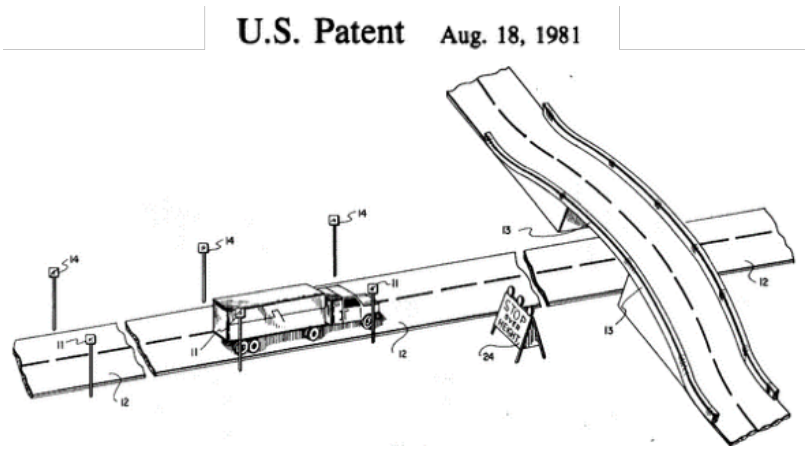


Fig. 14 First over-height vehicle detection and warning system, US patent in 1981

State	Manufacturer	EWDS used	System impact	Initial cost
Kansas	Elwood	Laser	Reduction	$500 + labour
Iowa	In house	Chains	Slight reduction	N/A
New York	In house	Headache bar	Slight reduction	N/A
Oregon	IRD	Laser	Reduction	$32K
Idaho	IRD	Laser	Reduction	$65K
Pennsylvania	IRD	Laser	Reduction	Unavailable
Florida	In house	Light	Reduction	Unavailable
Louisiana	IRD	Laser	Reduction	Unavailable
Mississippi	Unavailable	2 EWDS	Slight reduction	Unavailable
Maryland	Unavailable	Light	Reduction	$50K
California	IRD/Trigg	Laser	Reduction	$10 – 20K + Labour

Table 3 US Department of Transportation early warning detection system usage (Mattingly, 2003).

If the light beams are 'momentarily interrupted or broken', the successive pairs of light sources and sensors are activated. A signal is then sent to the control station which in turn activates a visible, flashing, electric sign indicating that the approaching vehicle is too high to clear the obstruction, and further warns the driver of the vehicle to stop or to exit from the roadway. A message concerning the OHV can be transmitted to the highway authorities simultaneously. A mechanical sensor can be located on the overhead structure, with an associated camera to take a picture of the vehicle if the driver fails to stop, causing an OHVS. A collision report can also be transmitted to highway authorities. A study by Mattingly (2003) shows that 38% of the following departments of transportation (California, Louisiana, Pennsylvania, Idaho, Oregon, Kansas) are currently using EWDS (Table 3). Headache bars are portal frame sacrificial systems.

Another system uses a patented Z-Pattern red/infrared dual beam array with the ability to reject ambient light, which eliminates false over-height alarms (Figure 15). A fault detection and alert function also notifies authorities in the event of a power failure. Agrawal (2011) reports that the installed cost is in the range of $7,700–8,900 USD plus a maintenance cost of $50 USD per year, however, in the same study, Maine reports to have installed a similar system in the range of $150,000–200,000 USD. The reason behind this price disparity is not known.

According to comparisons of different over-height detection systems in the study by Agrawal (2011), base installation costs ranged from $15,000–$20,000 USD per unit per direction. Table 4 shows the overall performance, reliability and effectiveness of each of the systems (Agrawal, 2011). In a study by the Michigan Department of Transportation to evaluate the use of OHVDs and warning systems, the cost of an active detection and warning system is estimated to be $110,000 USD, and the estimated 3-year benefit to be between $609,000 and $674,000 USD (Cawley, 2002). It is unclear how the cost and savings calculations were determined.

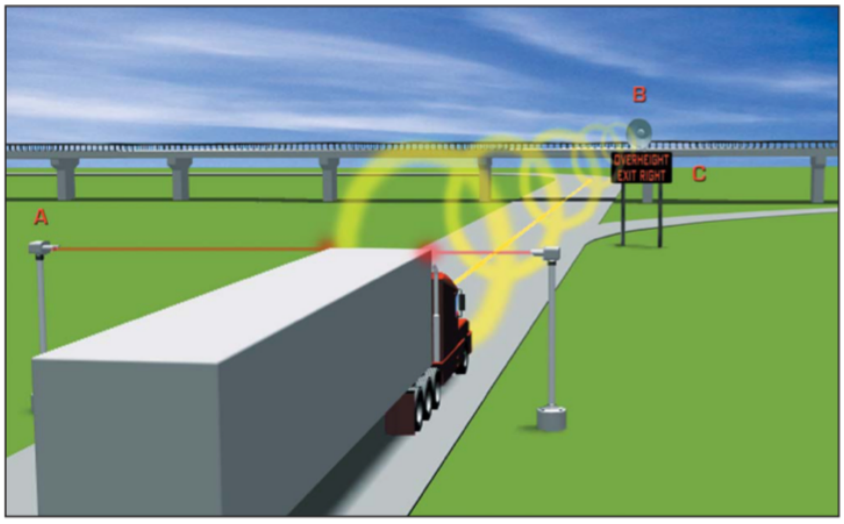


Fig. 15 TRIGG laser and sensor OHVDS. When an OHV breaks the sensor, the alarm bells and warning sign are activated. The sign message alerts drivers to use an alternative route

States	System type	Satisfaction with overall performance (out of 10)	Operational issues with system	False detection occurrences
Missouri	Z-Pattern	9	Lightning; vehicle strikes	None
Maryland	Optic	8	Insufficient space for installation	Sparse
Texas	Pipes on cable	8	Low speed/ volume roadways only	N/A
Hawaii	Infra-red/ LED/ IR	9	Difficult to access for maintenance	1/mth
Minnesota	Infra-red	8	Damaged by lightening	N/A
Maine	Z-Pattern	9	No	1/3 mths*
Alaska	Laser	1 (complications with false detections)	Many, too complex mechanisms, poor truck discrimination	Frequent (very sensitive)
Virginia	Dual beam	8.5	False detections caused by sun, bird activity	Frequent (caused by environmental factors)

Table 4 Agrawal's (2011) summary of effectiveness and user satisfaction of active systems
*Not frame-by-frame as in this work

1.3.4 Detection systems

Devices with the sole purpose of OHVS detection include sensors and accelerometers that are installed on the bridge structure to record any strikes. Strainstall developed a product called BridgeWatch, which provides structural monitoring data and information. It is a web-based system, providing real-time access to data. The detection sensor is used as a data acquisition system, collecting data at a single node for centralised processing. The sensor can notify authorities if an OHVS has occurred based on the multiple sensor inputs including strain gauges, accelerometers, temperature and displacement transduces, distributed across the deck and on the towers and piles of the bridge structure (Strainstall, 2017). An accelerometer can be used to parameterise a model of the structure: when damage occurs on the bridge structure, the parameters of the model change (Xu et al., 2004). The devices can be paired to a wireless network that can send the signal to a remote location for processing and decision-making.

Other electronic OHVS detection methods exist, such as fibre optic cables installed on parapets. The Mass Transit Railway (MTR) system in Hong Kong uses fibre optic cables to notify the station if a bridge has been struck. The technology uses a collision notification system to relay the message back to the control room (Byrne, 2009).

1.3.5 Reporting systems

Many OHVS accidents today are not reported, meaning bridge owners are left to remedy the damages caused by drivers. Bridge owners have been installing closed circuit television (CCTV) cameras near their structures, to capture the license plate of offending drivers (Martin & Mitchell, 2004, Byrne, 2009). CCTV technology can be used in combination with communications for incident detection and verification as well as weather and roadway conditions monitoring (Manjunathan, Albert & Deeter, 2005). The images can be sent wirelessly to a remote server using cellular telephone signals. Thakuriah & Geers (2013: p. 15) state that "The ability for bridge owners to have a complete picture of the environment is an essential ingredient for making intelligent operational decisions."

1.4 Conclusions and thesis overview

Figure 16 and Table 5 summarises the current state of practice under prevention, detection and reporting methods. The synthesis shows several advances on the passive end of the prevention methods however, methods remain sparse. The overall picture suggests the availability of many effective forms of detection and reporting systems for OHVS. Practitioners have access to readily available systems for the detecting and reporting aspects of OHVS management, but the devices alone will not prevent strikes; the main area of concern lies with prevention. OHVS still occur with high frequency, and OHVS prevention

systems (passive, sacrificial and active) on the market are often so expensive as to discourage widespread implementation.

Passive systems may be a 'quick fix' and cost effective, but they are not sufficiently effective overall, as evidenced by the common sight of scrape marks on the underside of bridges in the UK. Bridge owners aim to minimise the occurrences of OHVS and as a consequence, to minimise inspection, maintenance and repair costs. The need to develop an affordable yet reliable system is crucial to prevent future strikes posing risks to public civil infrastructure. The system should be affordable for the average low bridge, and not only targeted to specific problematic structures.

Existing EWDS are the most accurate warning systems, but they are not cost-effective due to their significant physical infrastructure requirements. Cost considerations drastically limit their adoption and suitability. Crash beams is one popular method used by department of transportations to prevent damage from occurring to structures. The limitations of the method are permit approvals and cost. The permit approvals may take months to obtain and the cost of constructing the beams are expensive to implement. The infra-red/laser beam has high-precision and accuracy; however, the cost of the system has prohibited widespread implementation. The LaserVision method is current under research and development at the Sydney Tunnel Bridge in Australia. The system is a cutting-edge advancement in this area of research however, the system is expensive, highly intricate and requires professional monitoring. This system is designed for the severe and critical case. The GIS, off-the-shelf alternative is a handy tool for truck drivers to pair with existing on-board navigation systems. Although cheap and portable, Transport of Scotland has advised lorry drivers to discontinue usage due to its often, out-of-date information. The system requires frequent updates and relying on the devices alone will not prevent strikes from occurring. Cameras are inexpensive in comparison to the sensor and laser systems. If the laser and sensor are effective at detecting OHV but the only barrier is cost, then the answer is to create a prototype that can behave like the sensor and laser at a significantly lower cost. The ideal system should be cost-efficient, robust to false detections in variable weather conditions while operating in real-time.

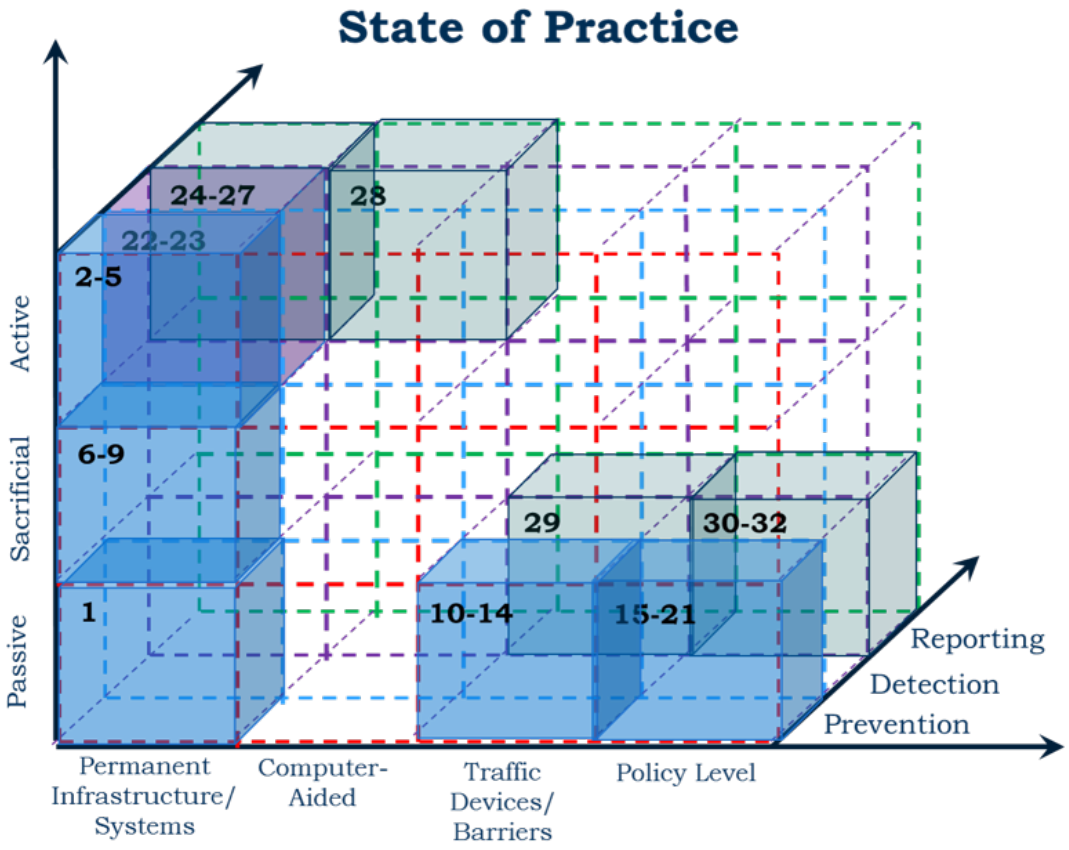


Fig. 16 Synthesis of current state of practice for bridge strike management

Prevention	Detection	Reporting
1. GPS: P	22. Piezoelectric: A	24. GSM: A
2. Lasers: A	23. Accelerometers: A	25. Remote Monitoring: A
3. Sensors: A		26. SMS: A
4. Infra-red: A		27. Email, broadband/server: A
5. Fibre-optics: A		28. CCTV camera: A
6. Sacrificial beams: S		29. Static signage: P
7. Portal Frames: S		30. Prevention groups: P
8. Metal strips/bells: S		31. Emergency service line: P
9. Speed bumps: S		32. Bridge identification number: P
10. Static signage: P		
11. Advance warning signs: P		
12. Variable message sign: P		
13. Flashing sign: P		
14. Rumble strips: P		
15. Driver education awareness: P		
17. HVG license test: P		
18. Good Practice Manuals: P		
19. Policies: P		
20. Fines/Fees: P		
21. Maps: P		

Table 5 Summary of current state of practice categorised under passive (P), sacrificial (S), active (A) methods.

New EWDS are needed that can bring the cost down by at least an order of magnitude to make them attractive to infrastructure owners. Therefore, this thesis presents a new solution for OHVD. The remainder of this thesis is structured as follows: Chapter 2 discusses the current state of research of OHVS technology, including gaps in research. Chapter 3 presents the work during a London Underground OHVS pilot project. This focused on researching problematic bridges within London, UK and installing a prototype set-up to test the current state-of-the-art research method. This chapter concludes with lessons learned and the objectives and research questions for the present thesis. Chapter 4 describes the proposed solution, hypothesis and contributions. Chapters 5, 6, & 7 test and Chapter 8 validates the solution in its constitute parts and as a unit. Chapter 9 concludes with contributions and future works.

Chapter 2

Current State of Research

2.1 Current state of research

Current research may be classified into two categories: computer vision methods, and sensor / laser methods. The next sections cover vision- and sensor- / laser-based methods in some detail. The chapter aims to synthesize the current state of research in OHVD. The chapter focusses on the latest computer vision, sensing and related methods for OH vehicle detection from feature extractions, height estimations to stochastic approximations. The capability for intelligent transportation systems to detect and track moving objects using vision-based systems is still challenging. However, with ever-increasing computational speed, applications of vision technology are increasingly viable. This next section explores camera calibration methods that are ideal for road traffic scenes. Next, the section explores related computer vision-based methods for object recognition and motion detection. These same concepts can be applied to the OH scenario for vehicle detection, tracking and classification.

2.1.1 Computer vision methods

Khorramshahi, Behrad, & Kanhere (2008) presents a new method for OH vehicle detection in low headroom streets using a digital video camera. In this method, Kanade-Lucas-Tomasi (KLT) and blob extraction algorithms are used to extract features for tracking. The OH system uses a cubic detection zone to obtain a vertical projection of a feature point on the road using blobs in 2D coordinates.

A cubic detection zone is constructed in the camera frame and marked by the user to indicate vehicle limits for varying sizes of trucks and buses. A blob detection method is used to detect regions in an image that differ in properties, such as brightness or colour compared to the surrounding region. The camera is located in a unique position to form a detection box that separates OH features from non-OH features. As OH vehicles pass through the detection box, OH features are tracked along the frame and segmented from the rest of the frame. Limitations of this study were that height estimations of feature points are not accurate in world coordinate systems when occlusion occurs. Occlusions occur

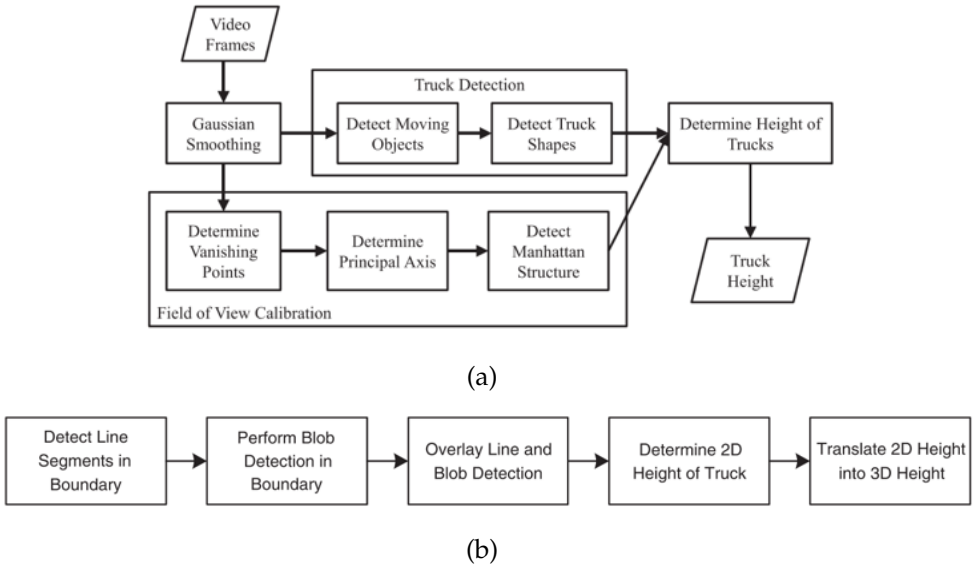


Fig. 17 (a) Overall framework of the over-height truck detection system and (b) height-determination workflow (Dai et al., 2015)

when part of the vehicle is blocked by another vehicle. Blob detection reliably extracts the silhouette (foreground) of vehicles (Danker & Rosenfeld, 1981; Hinz, 2005; Shneier, 1983). The foreground extracted can contain incorrect height estimation when one vehicle is occluding another in two or more lanes of traffic (Khorramshahi, Behrad, & Kanhere 2008: 289).

Shao, Zhou, & Chellappa (2010) measured object heights from video sequences in order to track moving objects (blob-based and feature point-based). The authors applied the single view metrology algorithm to each frame, using the least median of squares method as the cost function and the Robbins-Monro stochastic approximation as the optimisation algorithm. Single view metrology describes how 3D affine measurements may be computed from a single perspective view of a scene given minimal geometric information (Criminisi, Reid, Zisserman (1999). The Robbins-Monro stochastic approximation uses recursive update rules that can be used to solve optimisation problems by finding zeroes or extrema of functions which cannot be computed directly, but only estimated via noisy observations. Motion trajectory lines extracted from tracking results were prone to noise corruption. To deal with this, pre-specified thresholds were used to give approximate height measurements of the moving objects.

A recent contribution to height detection research is the work done by Dai et al. (2015), using line detection and blob tracking to locate the upper and lower points of a truck in pixel coordinates. These 2D coordinates are translated into 3D world coordinates to provide an approximation of the truck heights. The overall framework is presented in Figure 17a and height-determination workflow in Figure 17b. The Haar Cascade image training is supplied with images labelled truck or non-truck. The Haar Cascade method uses digital

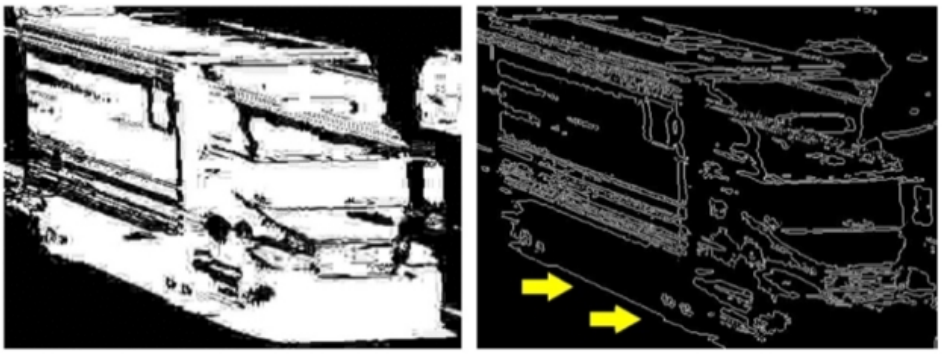


Fig. 18 Bottom boundary determination using (a) blob detection and (b) edge detection of the blobs (Dai et al., 2015).

image features used in object recognition to detect features on a set of training images and classifies the positive and negative images. Finding the top boundary of a given truck, the method uses three orthogonal axes and their vanishing points to find the principal axis. The Manhattan Structure method is widely used to extract vanishing points from visual scenes (Coughlan & Yuille, 2003; Furukawa et al., 2009; Mirzaei & Roumeliotis, 2011; Olufs & Vincze, 2011a; 2011b). To find the bottom boundary of the truck, blob detection and the Canny Edge detector were used (Figure 18).

The top boundary line is divided into n equal segments, and from each point, line scanning is performed to search for the intersection with the bottom boundary. The line whose inclination is the closest (or the most parallel) to the top boundary line in the real-world coordinate system is selected as the correct part of the bottom line. This research demonstrated promising initial results. However, major refinements are still required to mitigate the problems of occlusions, shadows and inaccurate line detection. Other limitations of the method include the use of Canny Edge Detection on the blob detection foreground mask.

If the edges that are extracted are not long enough, the Canny Edge Detector is unable to extract a continuous line (Marr & Hildreth, 1980). Therefore, if the blob detection encounters noise, edge detection will select false bottom reference lines, thereby producing incorrect height measurements (Danker & Rosenfeld, 1981; Hinz, 2005; Shneier, 1983).

The same concept applies to line segment detection: a noisy line segment may be generated from a road marking due to the limitations of the blob detection algorithm, and this in turn will contribute to falsely selected reference points (Von Gioi et al., 2010). Inaccuracy of height detection, then, is in general due to incorrect top and bottom lines selected to determine the height of the vehicles. Figure 19 shows video captured in rainy weather conditions. In severe rainy conditions, the camera lens may experience rain drops that may affect the detection process. Depending on road geometry and camera angle, trucks may be occluded in the field of view, in which scenario the proposed method suffers from false negatives: OH trucks are therefore missed.

Fig. 19 Video captured in rainy weather conditions (Dai et al., 2015).

Lastly, vehicle shadows reduce reliability and accuracy in OH detection performance at the pixel level, resulting in incorrect height measurements. Although acceptable preliminary results were achieved with pre-recorded data and manual measurement of trucks, the method is theoretically well suited for boxy vehicles with obvious boundaries (top and bottom) for discrimination and unsuited for irregularly shaped vehicles on the roadway.

2.1.2 Sensor and laser methods

In this section, three sensor / laser methods are reviewed. The key contribution of each study pertains to a different aspect of the holistic OHVS approach. Song, Olmi, & Gu (2007) developed an OHVS collision detection and evaluation system using piezoceramic transducers. Piezoceramic transducers were used for bridge impact detection and bridge health monitoring. An electric circuit is designed to detect the impact and activate a digital camera to photograph the offending vehicle as it collides with the concrete bridge. The proposed system thus has the simultaneous functionality of sending a triggering signal to capture an image of the offending vehicle, estimating the impact level and evaluating the damage level of the bridge. The system has shown potential to provide monitoring and accident notification.

Massoud (2013) presents a different approach for an OHVD system. The design encompasses mechanical, optical, and image processing methods. A laser & light dependence resister (LLDR) sensor is used in conjunction with a camera to transmit number plates of OHV to traffic administration bodies. The experiment can detect and estimate the height of trucks. Urazghildiiev et al. (2002) presented a novel vehicle classification system which used microwave (MW) radar measurement to obtain vehicle height profiles, in order to classify vehicles by height. The MW radar measurement provided high classification accuracy and performed well under most weather and illumination conditions. However, the sensor demands nontrivial computational capacity, and real-time operational performance is not reported in the paper. Follow-up research is presented in Urazghildiiev et al. (2007), which reported vehicle profiles (height / length) obtained by the MW radar sensor. Field trials were conducted using a spread-spectrum MW radar sensor showing high immunity to

Fig. 20 Irregularly shaped over-height vehicles.

vibration, adverse weather and illumination conditions. Although the MW sensor is ideal for determining the maximum height of a box-shaped truck (since the shape of the container is consistent), it is not reliable for capturing the maximum height of irregularly shaped vehicles.

Vehicles such as a flat-bed truck carrying an excavator or a truck with an extra-high muffler (see Figure 20) is less likely aligned with the sensor to indicate the highest point of the vehicle. Other limitations include the MW sensor being restricted to a single lane, and the high estimation errors when reflections from the vehicle surface are missed due to insufficient capacity of sensor. Both these limitations resulted in decreased classification accuracy.

The most recent advancement in laser-based method is presented by Singhal (2015) using LADAR (laser detection and ranging) technology installed on the face of the bridge overhead, referred to as LaRa in the approach. The method is unique, but the major flaw in this system is the sensor location (see Figure 21). If a bridge strike were to occur, the system could suffer significant damage and/or total destruction if the bridge were to collapse.

2.2 Related methods in the literature

Other related methods exist in the literature that have not been tested specifically on OHVS, but could enable effective solutions. They include the use of dynamic trip wires, background subtraction with Kalman filters, and perspective projection using a static camera.

2.2.1 Dynamic trip wires

Dynamic trip wires (DTW) can be used for motion detection applications such as pedestrian and vehicle counting, surveillance security applications and traffic monitoring systems. A DTW is a form of motion detection consisting simply of a line made up of pixels drawn on the image to trigger some process (Hardik, Shah, & Raval, 2013). The action is triggered when an object appears in or disappears from a specific region of interest (Haering, Venetianer, &

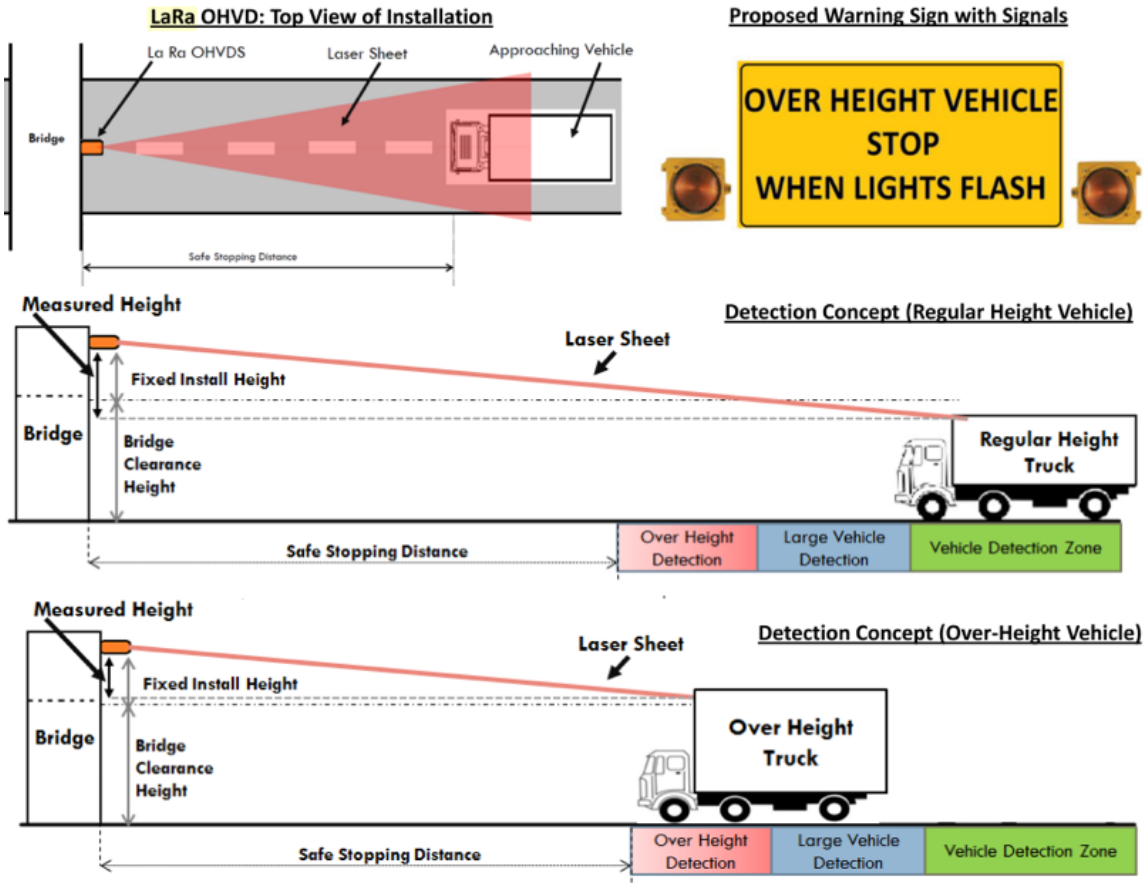


Fig. 21 Laser Ranging (LaRa) over-height vehicle detection method installed on face of bridge structure (Singhal, 2015).

Fig. 22 (a) Vertical trip wire shown in purple to trigger the alert of the alarm from vehicles travelling from right to left on the image; (b) Two dynamic trip wires (purple and cyan) to trigger an alarm when vehicles pass in either direction.

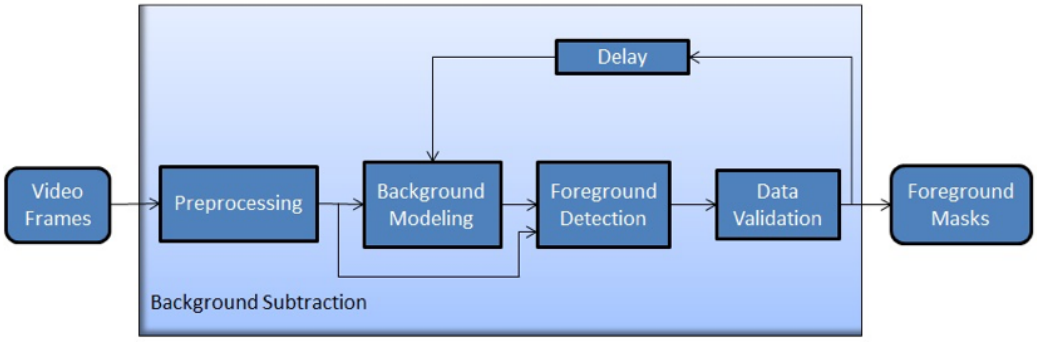


Fig. 23 Flow diagram of a generic background subtraction algorithm (Cheung & Kamath, 2005).

Lipton, 2008). The direction of tripping can be specified by the user and can be drawn in any direction as shown in Figure 22 by Jog & Halbe (2013).

DTW are useful in scenarios where there are two directions of traffic. Since DTW can detect objects entering from a specific direction, DTW can neglect traffic travelling in the direction which is not of interest, therefore this approach may save on computational costs including processing and power.

2.2.2 Background subtraction

Identifying moving objects from video frames is a fundamental task in computer-vision applications such as video surveillance, traffic monitoring and analysis and vehicle detection and tracking (Cohen & Medioni, 1999, Betke, Haritaoglu & Davis, 2000, Niu & Jiang, 2008; Yoo & Park, 2008, Shao, Zhou & Chellappa, 2010, Friberg, 2013).

A common approach is to perform background subtraction, which isolates the foreground image of moving objects (here, moving vehicles) from the portion of a video frame that differs significantly from a background model (Figure 23). Pixels in the current frame that deviate significantly from the background are considered to be moving objects.

The background subtraction method is robust when dealing with changes in illumination and using an adaptive median filtering approach can be effective against detecting non-stationary background objects such as rain, snow, falling leaves, and shadows cast by moving objects (Cheung & Kamath, 2004; 2005). A non-parametric model for background subtraction developed by Elgammal, Harwood, & Davis (2000) can handle situations where the background of the scene is cluttered and not completely static. The model estimates the probability of observing pixel intensity values for each pixel. The model adapts quickly to changes in the scene, which enables very sensitive detection of moving targets. The model runs in real-time and achieves very sensitive detection with 2% of false alarm (Elgammal, Harwood, & Davis, 2000: Pg. 16).

In conjunction with the background subtraction approach, Kalman filters can be an effective approach for OHVD as it functions well in free-flowing traffic (Coifman et al.,

1998). To accommodate for weather and light conditions, the Kalman filter-based adaptive method is widely used in computer vision to allow for robust background estimation under such variation (Kim et al., 2005, Trivedi, Gandhi, & Huang, 2005, Stauffer & Grimson, 1999; 2000, Wojek et al., 2010). Kalman filters represent an effective technique for tracking and can automatically adapt to changes in lighting and weather conditions (Cheung & Kamath, 2004) showing promise in OHVD for outdoor traffic scenes.

2.2.3 Perspective projection

Perspective projection can be used in the static camera scenario. The centre of the projection emits imaginary rays that intersect with the projection plane (Yang, Chen, & Beheshti, 2005). Let us assume there is an imaginary plane (made up of ray(s)) that is positioned above the road plane at x height. The x height is considered to be the height of the low clearance bridge where this imaginary plane lives. Below this imaginary plane, this is where non OHV's are present and above this line, vice versa.

The imaginary plane is determined using the pinhole camera model which defines the geometric relationship between a 3D point and its 2D corresponding projection onto the image plane (Morvan, 2009). Perspective projection is the geometric mapping from 3D to 2D using the pinhole camera model. A camera can be calibrated using the perspective transformation model as in eqns. 1 and 2:

$$s\,m' = A[R \mid t]M' \tag{1}$$

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2}$$

- X, Y, Z are the coordinates of a 3D point in the world coordinate space
- u, v are the coordinates of the projection points in pixels
- c_x, c_y is a principal point at the image canter in pixels
- f_x, f_y are the focal lengths expressed in pixel units
- M' gives the 3D position of the camera (6 degrees of freedom).

The extrinsic parameters $[R \mid t]$ give the position of the camera in terms of its rotation and translation. The intrinsic parameters A are independent of the camera position and subsume:

- f is the focal length in meters

- s_x, s_y is the pixel width and height in meters
- c_x, c_y is the image centre in pixels
- K is the lens distortion

The image rays can be used to find the position of the imaginary plane in a traffic scene using perspective projection. However, lens distortion must also be taken into account. Most camera lenses suffer from radial lens distortion which causes straight lines to be mapped as curved lines therefore affecting the 3D position of the imaginary plane with respect to the road plane. Prescott & McLean (1997) used the radial distortion parameters to iteratively estimate a distortion model, warping the distorted image to the correct observed distortion, and evaluating the result to determine if distortion had been adequately eliminated. The results of the method corrected images to within 0.4-pixel average deviation of the actual distortion in the images, reducing distortion by up to 98.6%. With such corrections, perspective projection is a promising method that can be used for the OHVD scenario mapping ray projections in the traffic scene.

2.2.4 Camera calibration methods

When setting up a camera for OHVD, calibrating the camera is completed first. A commonly used calibration method is that of Tsai (1987). The method has been implemented in many calibration papers, including Ismail et al. (2013), Sturm & Maybank (1999) and Kanhere & Birchfield (2010). The model is based on a pinhole perspective projection model. The intrinsic camera parameters can be found via the camera hardware specs and the extrinsic parameters can be estimated from a non-coplanar set of world points where:

- f is the focal length of the camera
- k is the lens radial distortion coefficient
- u_0, v_0 is the principal points and the centre of radial lens distortion
- S_x is the scale factor to account for imperfections due to hardware
- D_x, D_y is the distance between adjacent sensor elements (fixed parameters)
- X_f, Y_f are the final pixel positions in the image
- R_x, R_y, R_z are the rotational angles for the transformation between the world and camera coordinates
- $r_{11}, r_{12}, r_{13}, \ldots$ are the rotation coefficients
- t_x, t_y, t_z are the translation components for the transformation between the world and camera coordinates

- X_w, Y_w, Z_w are the world coordinate points
- x_i, y_i are the undistorted image plane coordinates
- X_d, Y_d are the distorted image plane coordinates
- X_f, Y_f are the final undistorted points coordinates
- r is the transformation of distorted points in the image plane to the final points

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{3}$$

The transformation from 3D world position (X_w, Y_w, Z_w) to the image plane (X_i, Y_i, Z_i) is computed by undistorting the image plane (X_i, Y_i) co-ordinates as

$$x_i = f\frac{x_i}{z_i} \text{ and } y_i = f\frac{y_i}{z_i} \tag{4}$$

where the transformation from undistorted (X_u, Y_u) to distorted (X_d, Y_d) image coordinates is expressed as

$$x_u = x_d(1 + kr^2) \text{ and } y_u = y_d(1 + kr^2) x_u = x_d(1 + kr^2) \text{ and } y_u = y_d(1 + kr^2) \tag{5}$$

where $r = \sqrt{(x_d)^2 + (y_d)^2}$. The transformation from distorted co-ordinates in the image plane (X_d, Y_d) to the final image co-ordinates (X_f, Y_f) are:

$$x_f = \frac{s_x x_d}{d_x} + u_0 \text{ and } y_f = \frac{y_d}{d_y} + v_0 x_f = \frac{s_x x_d}{d_x} + u_0 \text{ and } y_f = \frac{y_d}{d_y} + v_0 \tag{6}$$

as the final undistorted pixel position in the image. This method is suitable for calibration of the OHVD camera, and is assumed in what follows.

2.2.5 Optical flow (motion)

Yoo and Park (2008) presents a novel approach for detecting moving objects in the camera view, using what they term the Earth Mover's Distance to find motion patterns in a given region. The algorithm subtracts two consecutive frames from each other and assigns motion blocks to detect regions with movement; it shows robustness to local illumination changes.

Similarly, Mittal and Paragios (2004) present a patented technique for modelling dynamic scenes using kernel-based multivariate density estimation for motion detection. The technique performs well under adverse weather conditions and rigorous background motion (such as moving trees and bushes); the algorithm is able to minimise background noise and therefore represents a good foundation for OHVD.

Niu & Jiang (2008) present an improved adaptive background subtraction detection method using a Gaussian mixture model to minimise shadow interference from moving objects. The method shows robustness to various illumination changes caused by vehicle shadows and lighting changes. Adaptive background subtraction is thus also promising for OHVD under variable weather conditions.

2.2.6 Feature detection, tracking & classification

Zheng & Chellappa (1995), Yao & Chellappa (1994), Tomasi and Kanade (1992) and Chetverikov & Verestói (1999) have all presented effective methods to detect moving objects using feature-based detection, tracking & classification. Tomasi and Kanade present a widely used method using factorisation to track the motion of features in an image stream. The method uses the size of eigenvalues to detect corners and regions with high spatial frequency content and intensity variance. The method compares past and present fixed-sized feature windows by taking the sum of the squared intensity differences over the windows and finding the displacement of one frame to the next using texture-rich pixels. The method shows robustness to occlusions and noisy images, which makes it ideal for OHVD and tracking.

Feature detection and tracking is a crucial step in preventing false positive detections for OHVD.OH vehicle detection. Vision-based methods show promise for OHVS, but despite their cost efficiency, asset owners are not universally convinced that vision-based systems are suitable to handle the vigorous outdoor conditions while maintaining performance accuracy. Further testing is required to achieve and demonstrate the effectiveness and value of the approach. If the accuracy target of ±5.00 cm can be achieved, a vision-based system (paired with complimentary detecting and reporting tools) could provide a holistic solution to the OHVS problem.

2.3 Summary of current state of research

Figure 25 and Table 6 synthesises the current states of research within the three categories of permanent infrastructure, computer-aided systems and active sensor systems. The current state of research shows a concentration in the area of prevention, while detection and reporting is already quite efficient. As for computer vision and machine learning techniques, these are still a developing area of research as computer processing units continue to increase in speed and affordability.

The methods to detect OH trucks presented by Dai et al., (2015), Massoud (2013), Shao, Zhou, & Chellappa (2010) and Urazghildiiev et al., (2002) used algorithms to extract the height of vehicles, whereas Khorramshahi, Behrad, & Kanhere (2008) used a cubic detection zone to obtain vertical projections of feature points on the road using blobs in 2D coordinates focusing on the region that is OH. In Figure 24, the image shows the result of marking procedure for a truck with known dimensions. The cubic approach is a promising start point

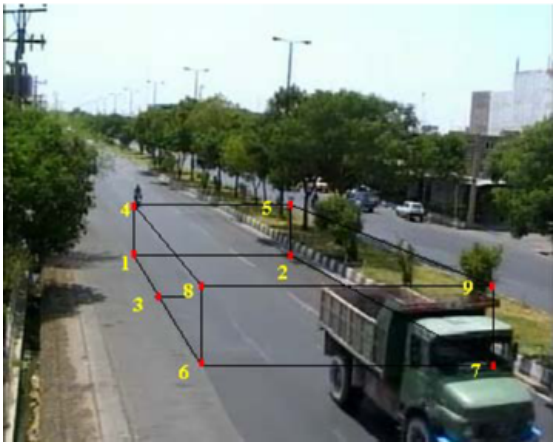


Fig. 24 Over-height vehicle detection using cubic detection zone to track features within image (Khorramshahi, Behrad, & Kanhere, 2008).

for OH detection. The approach allows for the detection of vehicles regardless of shape and size, while extending beyond only box-shaped vehicles. Although the current computer vision methods have not been tested in real-time, from a theoretical perspective there are flaws in each of the approaches. The Khorramshahi, Behrad, & Kanhere (2008) method is adversely affected by occlusions in the scene, therefore contributing to false positive and negative detections of OHV. Methods Shao, Zhou, & Chellappa (2010) and Massoud (2013) provide approximate height estimations of the vehicle. Bridge owners expect a certain degree of accuracy therefore; height approximations of vehicles is not acceptable in the case of OHV detection. The Urazghildiiev et al. (2002) method uses a microwave sensor to take a cross-section of vehicles passing under the sensor. The tallest part of the vehicle must pass directly under the sensor; therefore, the method prone to missing irregularly shaped OHV. Lastly, the Dai et al. (2015) method cannot detect non-box-shaped vehicles; neither can it process data when occlusions occur or under variable illumination conditions. Under these circumstances, the methods cannot fulfil their theoretical role hence; implementation in real-time is not required.

For current OHVS prevention systems, the biggest issue for bridge owners is affordability and reliability, without compromise on the accuracy and performance of such a system. Despite their potential significance, vision-based methods have received little attention as a potential solution to the OHVS problem. There has been little research done in computer vision on OHV detection, and no vision-based system has been implemented.

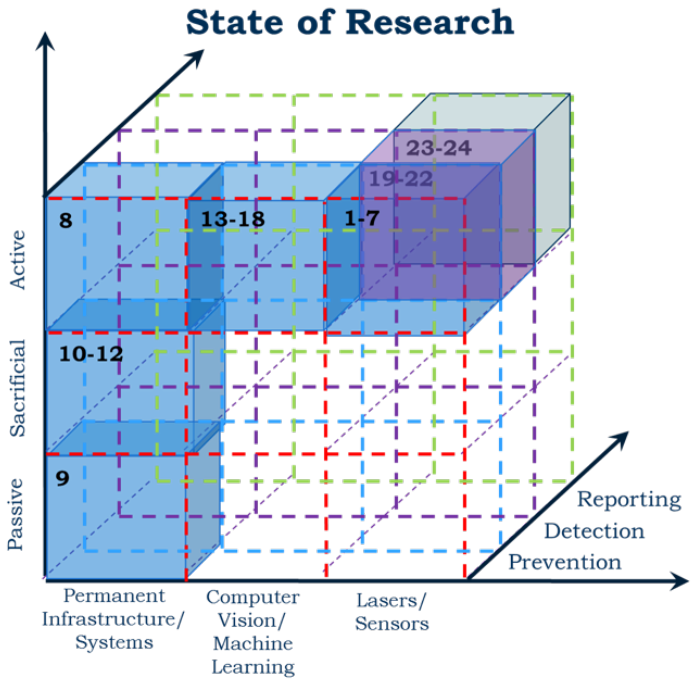


Fig. 25 Synthesis of current state of research under permanent infrastructure, computer-aided and active sensor systems.

Prevention	Detection	Reporting
1. Gupte et al. 2002	19. Xu et al. 2004	23. Song, Olmi & Gu 2007
2. Urazghildiiev et al. 2002/ 2007	20. Song, Olmi & Gu 2007	24. Massoud 2013
3. Song, Olmi & Gu 2007	21. Mehrani, Ayoub, Ayoub 2009	
4. Liu, Chen & Hasuer 2012	22. Massoud 2013	
5. Massoud 2013		
6. Sinfield 2010		
7. Singhal 2015		
8. Hanchey & Exley 1990		
9. Horberry et al. 2002		
10. Qiao, Yang & Mosallam 2004		
11. Sharma, Hurlebaus & Gardoni 2008 / 2012		
12. Yang & Qiao 2010		
13. Khorramshahi, Behrad & Kanhere 2008		
14. Zhang, Avery & Wang 2007		
15. Shao, Zhou & Chellappa 2010		
16. Sandidge 2012		
17. Park et al. 2013		
18. Dai et al. 2015		

Table 6 Summary of research conducted in over-height vehicle strike prevention under the categories of prevention, detection and reporting.

The objective of the study is to develop a vision-based system using a single camera to detect OHV while minimising the number of false positive detections. Positive results in this endeavour will feed direct warnings to drivers as well as reporting in the case of strikes. As remarked above, cameras are roughly an order of magnitude less inexpensive than typical sensor / laser systems, and if the vision-based system developed here can yield an accuracy of 5 cm or less, it will represent a desirable solution for asset owners.

2.4 Hypothesis and aims

The research hypothesises that installing a vision-based camera at the height of the low bridge can accurately detect OHV and warn drivers for the prevention of bridge and tunnel strikes. The overall aims are to develop and test a novel vision-based system for detection of OH vehicles. The approach intends to reduce the number of overhead strikes, while prolonging the life of bridge and tunnel infrastructure and minimising the cost of inspections, maintenance and repairs. The subsequent aims of the research are to:

- **Aim 1:** Determine the calibration process, scene configuration and setup of the camera including optics, placement and orientation of an OH camera.
- **Aim 2:** Determine the optimised image initialisation parameters to accurate detect OHV.
- **Aim 3:** Minimise the number of false positive detections of OHV.
- **Aim 4:** Validate the overall performance of the OHVD system using the optimised parameters.

Chapter 3

London Underground case study: Bridge strikes

3.1 Introduction

A feasibility study was conducted in London (UK) using a computer vision approach first proposed by researchers at the Georgia Institute of Technology (Park et al., 2013; Sandidge, 2012). The study tested the viability of a digital video system combined with the aforementioned computer vision algorithms to determine the height of approaching vehicles using a single camera. The project aimed **to test and validate the over-height algorithm under real-world conditions to determine its accuracy.**

A preliminary system was applied to a set of London Underground (henceforth LU) bridges that encountered frequent OHVS. The system monitored traffic on a single approach, with London Underground requiring an accuracy within 5 cm of the actual height of the vehicle. The feasibility study was essential to remedy gaps in the existing literature and report the results of real-time processing. The LU Project began on October 6th, 2013 and ended on April 6th, 2014 (6 months). The project activities are presented in Table 7. The beginning of the project consists of the initialisation and planning stages. An end-user requirement was conducted to determine the features and functionalities of the system. Necessary hardware and equipment was purchased at this stage. Preliminary data was collected and compared to existing data vehicle height data where vehicle heights are known. At this stage, the preliminary data was used to test and debug any errors in the system. A full-scale validation was performed once the preliminary validation stage is completed.

Activities	Description
Research/planning	Research methodology and experiment layout; Planning, plus acquisition of the necessary hardware (camera, lens & communication device).
Data collection	Prototype testing on selected bridges experiencing frequent OHVS (Preliminary and validation data collection).
Preliminary validation stage	Analysis of height accuracy results against using both line video and manually collected validation data.
Full-scale validation stage	Validation of algorithm, hardware, and accuracy measurement of vehicle heights.
Report preparation	Reporting of results to London Underground for further evaluation.

Table 7 Activities of Bridge Strike project.

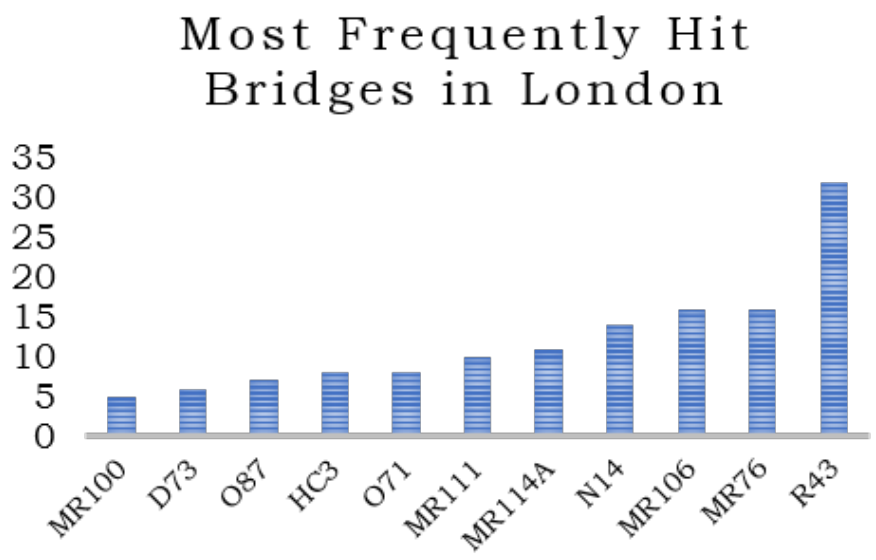


Fig. 26 Bridges prone to strikes in London in 2013.

3.2 Background

In the UK, a low bridge is considered one with a clearance of 5.03 m (16′ 6″) or lower. Annual OHVS data for the period of 2003-2013 on LU networks indicate a peak in OHVS occurrences in 2010. In practice, especially in more remote areas of the network, OHVS may go unreported, meaning potentially unsafe bridges continue to operate. In West London, South Ruislip is the most frequently struck bridge.

When strikes occur, LU must take extra measures to alert drivers to the reduced clearance, for which they presently usually employ static warning signs. Despite such warnings, bridges continue to be struck on the LU network. The majority of damage generates costly repairs to bridge infrastructure and causes delays on the transportation network. The bridges most hit is depicted in Figure 26 (R43, MR76).

The cost associated with such strikes has increased the urgency of exploring better methods of prevention. The overall goal is to reduce OHVS occurrence and consequently to minimise inspection, maintenance and repair costs. LU has therefore sought an early warning system to both warn drivers of impending OHVS and to send notification if strikes occur.

3.2.1 Case study locations

A number of test sites were considered for the feasibility study; these are listed in Table 8. R43, located at South Ruislip already has a laser detection system, which, however, is not functioning correctly at the time of this study. R43 was most frequently hit due to its location profile. The bridge was located in a highly built-up area containing many connecting (artery)

Bridge ID	Location	Height	Posted Speed
R43	South Ruislip Station	3.6m (11′-9″)	40 mph
MR76	Moor Park to Rickmansworth	4.3m (14′-3″)	30 mph
N14	Fairlop Station	4.2m (14′-0″)	40 mph

Table 8 Site characteristics of frequently struck bridges in London, UK.

Fig. 27 R43, South Ruislip Station in west London. The over-height detection system is not working (hence sign posted in 2013).

roadways. Site inspection revealed many obvious signs of damage, including scrape marks on the underside of the bridge (as shown in Figure 27). However, R43 was not selected as the test site of the feasibility study, as there was no clear 1000-metre stretch of roadway along which the camera could be placed. Also, several local streets connected to the main roadway leading to the station. The system required a clear roadway of 1000 m with no connecting roadways, otherwise, it is possible that OH vehicles would be missed by the OH camera. According to the UK Department of Transport (2007), the stopping distance for cars on 30 – 50 mph roadways are between 23 - 53m for thinking and braking distances. For a lorry, the braking distances are much longer as the vehicle is much larger and heavier, so these figures are a minimum. In addition, the latency of the system requires ±500 m from the time of detection to time of warning. A safe turnaround distance should also be considered to allow for truck manoeuvring (if not possible to achieve off-site). The rule of thumb is to allow for an approximate distance of 1000 m from the bridge to the location of the camera to provide the system with sufficient time to detect and warn the driver and allow for the driver to react to the warning.

MR76 was considered next, located in Watford (Northwest London). The location of this bridge was ideal as there is a clear stretch of uninterrupted roadway and located in a rural area (see Figure 29). However, the camera and processing unit required access to power and no source was available at this location, which prevented further consideration of this site.

Fairlop Station Bridge (N14) in East London was chosen as the test location to mount the camera system. The camera was connected to a computer system, which processes the video stream and applied the OH detection algorithm (Figure 28). The total cost of installation is

Equipment and hardware	Unit cost
Axis outdoor camera	£1,140
Processing unit + 16GB Ram	£1,241
4 x 1TB hard disks	£264
Router	£25
Extras (outdoor utility box, cables)	£140
Total	£2,810

Table 9 Cost of the vision-based system.

Fig. 28 Camera setup installation at Fairlop Station in London, United Kingdom.

presented in Table 9. The system was installed at a height of 4.2m approximately 800m east of Fairlop Station angled at approximately 30° clockwise perpendicular to the road plane, facing traffic (these being optimal height and angle for OH detection). The camera was mounted on the south side of Forest Road, approximately 1000m from the bridge structure (Figure 30). The distance from the camera location to the proposed truck turn-around area is 620m, which would provide the truck driver sufficient time to react, brake and turn around. The camera was mounted on an existing streetlight pole and a truck turn-around area was designated.

3.3 Proposed approach

Permission was obtained from the Georgia Institute of Technology (GT) to allow Cambridge Engineering Department, Laing O'Rourke Centre, Construction Information Laboratory (CIT) to modify and use an existing code by Park et al. (2013), referred to as the *Park method*. The Park method uses a single camera to detect box-shaped trucks on the roadway and calculates its height using computer vision methods. A trigger is alerted if the truck is over a specific height. The overall framework is described in Section 2.1.1. The code was translated into the C# programming language, using the Microsoft Visual Studio .NET IDE, in order to handle real-time processing of video sequences; the translation is henceforth referred to as the *modified version*. The EmguCV Library, which is an open source .NET wrapper of the

Fig. 29 MR76, Moor Park to Rickmansworth (struck 16 times in 2013).

Fig. 30 N14, Fairlop Station located in East London, United Kingdom.

OpenCV image processing library was employed to allow access to OpenCV functionality within a C# framework.

After code translation, the algorithm was trained on representative images. The Park method used the Computer Vision System Toolbox for training (Haar Cascade method), whereas our modified version used EmguCV's cascade classifier. The Park method used 200 positive samples, whereas the modified version required a larger training set. MATLAB, used by the Park method, has a larger image training set than the EmguCV classifier, hence the lower number of training samples they required. The training set consisted of 450 positive samples, containing trucks, buses and full-size vans and 400 negative samples not containing the former. According to the OHVS data provided by London Underground (p.c.), OHVS has been caused by coaches, large vans, and lorries so these vehicles were included in the training sample. Since the samples all have similar 'box-shaped' features, this was an acceptable modification of training data. The Park method used training data consisting solely of box-shaped vehicles.

The next step involved calibrating the raw video sequence using reference points from the scene. Calibration was performed according to the Single View Metrology approach of Criminisi, Reid, & Zisserman (2000). The method uses vanishing lines on a reference plane and vanishing points from a reference direction up to a common scale factor. The utility pole is used as the reference object for calibration. The top and base point of the reference segment was manually set in the algorithm. Calibration was performed manually using MATLAB's built-in function to simultaneously calibrate the camera. There was no drawback to manually calibrating the code at start-up. This is a common process in camera calibration (Tsai, 1987, Kanhere & Birchfield, 2010; Shao, Zhou & Chellappa, 2010).

After training the sample images and calibrating the camera, the raw video sequences were processed. The MATLAB version used the built-in blob detection function, whereas the modified version used the EmguCV blob detection function to subtract the foreground of the image (in this case, the vehicles entering the field of view) from the static background image. The main difference is in the initialisation process. The MATLAB version required no blob initialisation, while setting of blob width and height is required in the modified version. When a truck is detected, a rectangular box is applied around the truck as shown in Figure 31. The Manhattan Structure is used to determine the vanishing points along the road, width and vertical plane (x, y, z planes respectively).

The Manhattan Structure is an assumption that states that the built environment were constructed on a Cartesian grid which enables us, from a single image, to determine the orientation of the relative object in the scene (Coughlan & Yuille, 2003). The algorithm detected the lines in the bounding boxes and classified it using the Manhattan Structure. The algorithm then determines the lane on which the vehicle is situated. This is required to find the truck plane which is along the width of the road and perpendicular to the road surface. The modified version used EmguCV's line segmentation detector to determine the top boundary of the truck and its blob detection method to determine the bottom boundary

Height validation sensor	Unit cost
Continuous laser sensor (height measurement)	£1,500
Data logger	£390
Total	£1,890

Table 10 Cost of height validation sensing equipment.

of the truck. Canny edge detection is used to find the longest continuous line at the bottom of each detected blob.

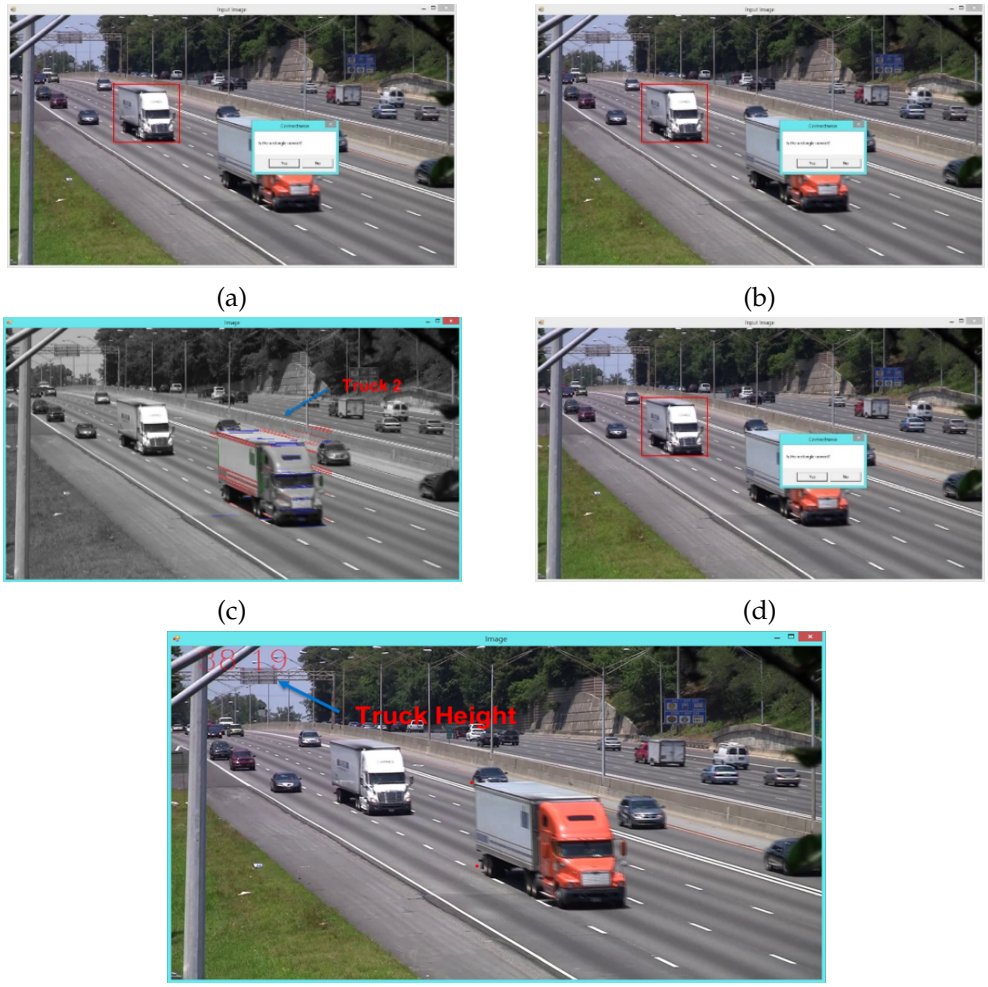
The Canny edge detection is a multi-stage algorithm that detects a wide range of edges in an image from extracting useful structural information from objects (Canny, 1986). Using this, the top and bottom boundaries of the truck were determined in order to yield the height of the truck in 2D pixel coordinates. Based on the reference dimensions, the 2D height calculated in the image frames was then converted to a real-world metric value. The Park method used line scanning method to search for lines using sub-segments which intersected with the bottom boundary. The line whose inclination is closest (or most parallel) to the top boundary line is chosen as the bottom line.

3.4 Experiment and results

Video was obtained from a single camera mounted 1000m upstream of the bridge structure (N14). Raw video footage was recorded using the Axis 1765-LE camera in London (UK). The camera has infrared capabilities suitable for night-time conditions. Figure 32 shows a validation sensor (TruSense) to validate the height of trucks. TruSense is a continuous distance measurement device with an accuracy of 2 cm. The sensor was mounted to a crane bucket 10 m above the lane of traffic (facing perpendicular to the road plane) to capture the height of the trucks as they passed underneath the laser beam. Simultaneously, the raw video footage was recorded (cost breakdown supplied in Table 10). The sensor was set to record at 100Hz (Table 11). Details of the sensor scanning interface can be found in Figure 33. Identical date-stamps were set on the sensor and camera to ensure that each truck in the video sequence was linked to the correct height measurements. Raw video footage was analysed in the laboratory.

Using the computer-vision method, 65 trucks were selected from the collected data to validate the accuracy of the method. Out of those 65 samples the algorithm analysed all objects that had a 'box-shaped' structure. This included HGVs, buses/coaches and light-goods vehicles. Although not all vehicles were considered OH, the purpose of the feasibility study was to determine how accurate the 3D detected image height is in comparison with ground truth validation (TruSense).

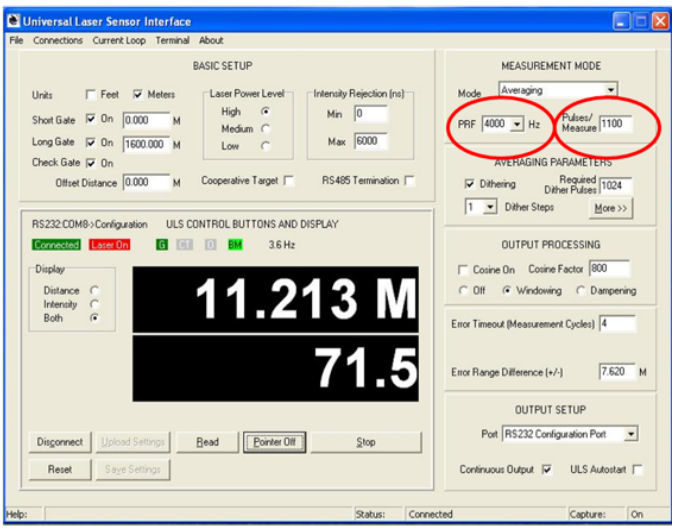
The summary presented in Table 12 show the results of the implementation. The results of the study showed a larger variation in the standard deviation from the average sample



(a) (b) (c) (d) (e)

Fig. 31 (a) The bounding box is detected around the Truck 1 and prompts the user to verify whether the rectangular box is correctly placed; (b) Applies the Manhattan Structure to the features within the bounding box for Truck 1; (c) Same process is performed in (b) for Truck 2; (d) Algorithm determines the top and bottom boundary of the vehicle and; (e) Displays the height of the truck in centimetres.

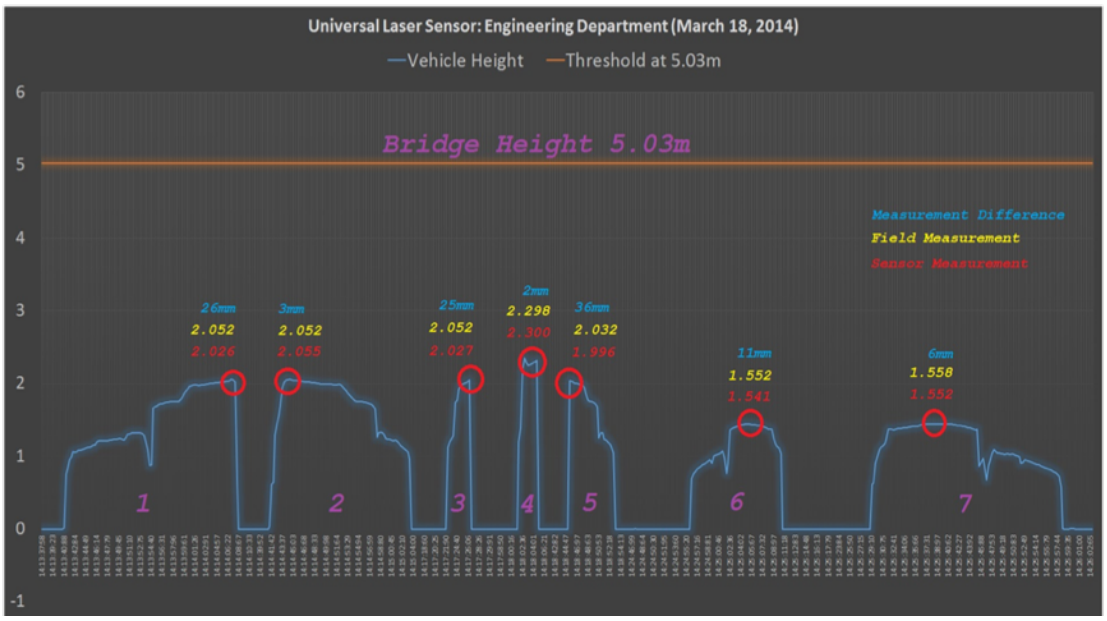
Fig. 32 Validation sensor paired with Acumen Instruments DataBridge logger.



(a)

Height of Sensor: 11.212 m		
Threshold: 5.03 m		
Time	Measured	Height
14:13:37:88	11.215	0.001
14:13:37:98	11.214	0.001
14:13:38:13	10.216	0.996
14:13:38:23	10.152	1.041
14:13:38:38	10.163	1.149
14:13:38:48	10.147	1.265
14:13:38:53	10.133	1.379
14:13:38:63	10.125	1.587
14:13:38:78	10.115	1.897
14:13:38:88	10.108	2.104
14:13:38:93	10.097	2.195
14:13:39:03	11.214	0.001
14:13:39:18	11.191	0.001
14:13:39:23	10.453	2.198
14:13:39:33	10.391	2.096
14:13:39:48	10.255	1.901
14:13:39:59	11.212	1.467
14:13:39:64	11.211	1.036
14:13:39:74	11.211	0.001

(b)



(c)

Fig. 33 (a) Laser sensor interface and (b) distance measured from the sensor to the top of the asphalt pavement; (c) Silhouettes of vehicles captured by the sensing device as they pass under the sensor.

Output Update Rate (Hz)*	PRF**	PPM***
1	1000	1000
10	2000	200
100	4000	40
1000	4000	4

Table 11 Height validation sensor output readings in Hertz.
*Output Update Rate = PRF/PPM per second
**PRF = Pulse Rate Frequency
***PPM = Pulse per Measurement

# of Trucks	# Measured Avg	3D Image Height Avg	True Heights Validation	Accuracy	Error Rate
65	61	2546	2786	91.05%	8.95%

Table 12 Results of the prototype.

height. The standard deviation is calculated based on the sum of the total 3D heights and true height validation. This dispersion is caused due to the large spread in height measurements i.e. heights of semi-trucks, buses and full-size vans; in this case, the standard deviation is rather insignificant. The 3D measured height is 2546 mm and true height validation is 2786 mm by calculating the averages of the 65 samples. The accuracy is calculated based on the average of the sum of the error rates for the 65 test samples. The error rate is 8.95% and determined based on the height difference from the 3D image height and true height validation

In Table 13, the results show the percentage of trucks detected with a truck height error of 5, 10, and 15 cm and under. For example, only 32% of detected trucks were within 5 cm of the true vehicle height. On the contrary, a higher percentage, 70.8% of vehicles were detected within 15 cm of the true vehicle height. In Figure 34, a graph is plotted of the sample set to indicate the height variation between the 3D detected image height and the true height measured with a laser sensor. The blue line represents the height of the bridge at 4.2 m. The data is sorted from the shortest to tallest height measurements. Out of the 65 samples, only one bus was considered to be OH i.e. > 4.2 m. The bus was headed to a nearby recreational centre and therefore, did not pass underneath the bridge.

5 cm and under	10 cm and under	15 cm and under
32.0% of trucks	61.5% of trucks	70.8% of trucks

Table 13 Results for truck height errors under 5 cm, 10 cm, and 15 cm ('under' refers to the height accuracy of the algorithm).

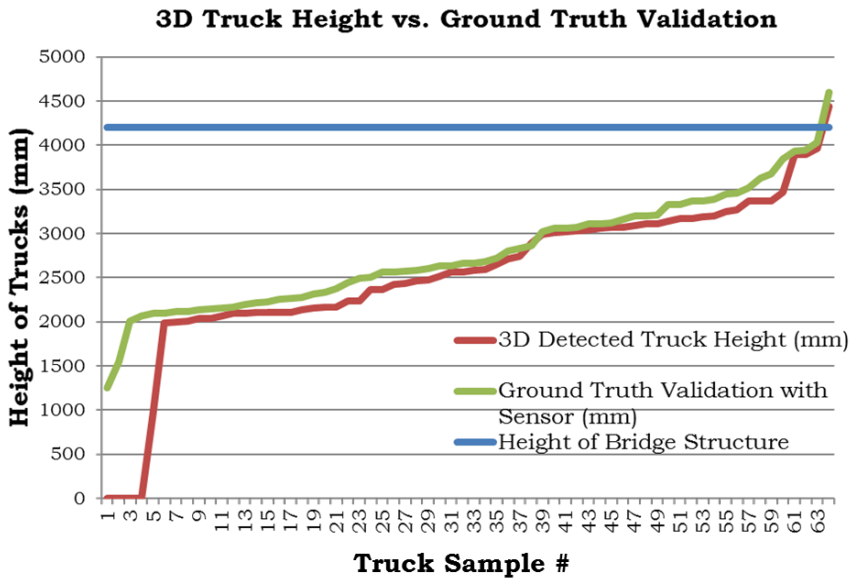


Fig. 34 Individual trucks ordered based on height measurements.

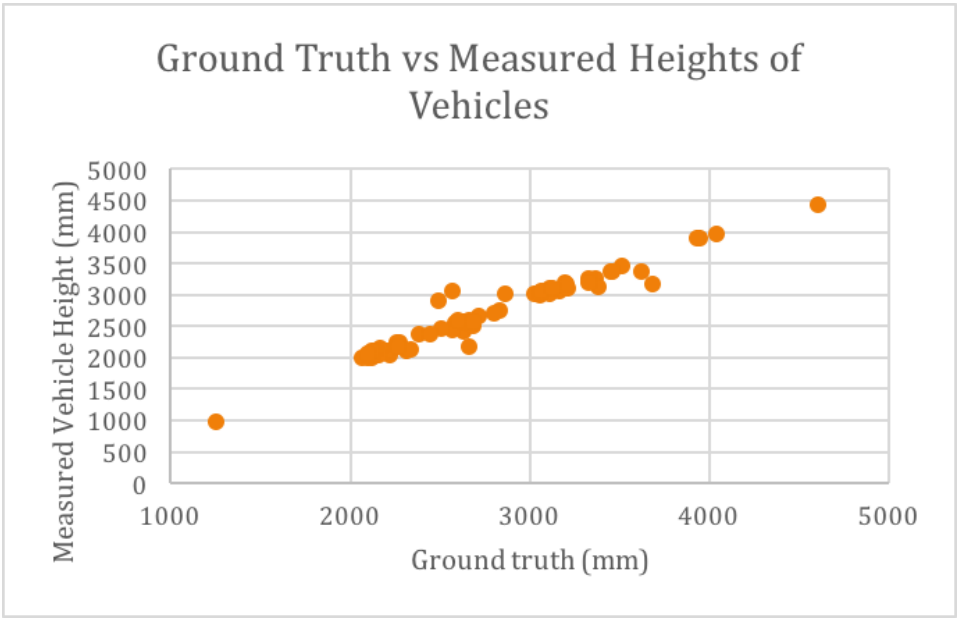


Fig. 35 Measured height of vehicles vs. ground truth (mm)

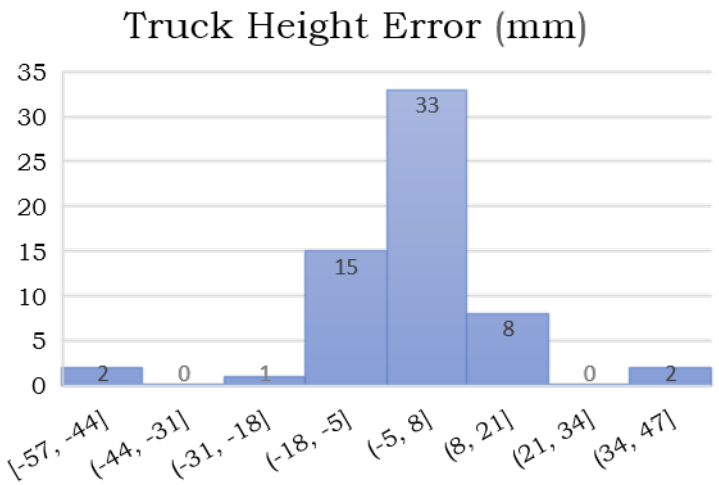


Fig. 36 Truck height error (mm)

Figure 35 shows a scatterplot of the ground truth data along the x-axis and measured vehicle height along the y-axis. Four of the vehicles were excluded from the data because they were not detected by the camera due to their non-box-shaped features. The data shows that there were no large discrepancies between the detected and measured with an average height error of 8.95%.

Figure 36 shows a histogram of the overall measured truck height error in mm. The ground truth is plotted at zero to indicate the base-line for comparison. The major discrepancies occurred at observations 6, 7, 18 and 58 showing that vehicles with box-shaped features such as dump trucks and flat beds can cause inaccurate height extractions as the top and bottom boundaries may vary due to the multiple surfaces. The algorithm may extract the cab height as opposed to the most upper and lower limits contributing to height extraction errors.

3.5 Discussion and conclusions

The results of the study are presented in Table 14. During the testing phase, there were four trucks (12, 29, 43 and 49) that were not measured by the algorithm. This is due to the truck detection algorithm, which may occasionally (6% in this case) miss a passing truck due to the truck not having 'box-shaped' features that were used for training (the detection algorithm is trained with 450 samples of trucks, buses and full-size vans). This is a limitation of the algorithm. The samples 6, 18, 58 measured a height taller than the actual height.

The feasibility study showed that the method worked as expected. By eliminating a systematic error of 8cm (i.e. increasing all measurements by 8 cm), the average truck height error is measured to be ± 8 cm. Figure 37 shows the truck height errors. 80% of the trucks were measured with an error less than ± 10 cm. Further improvements are required to

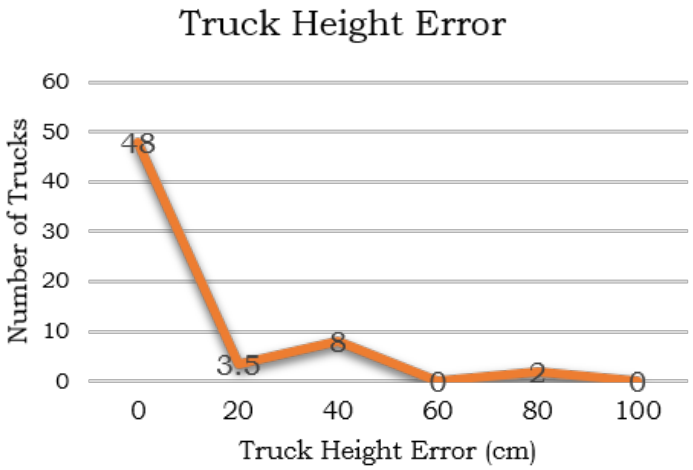


Fig. 37 Truck height error.

resolve certain niche issues to further improve the method. First, the method is limited to only box-shaped vehicles, up to two lanes of traffic, and daytime conditions. The method is occasionally prone to vehicle occlusions in the field of view which resulted vehicles being missed in the few cases observed where height traffic caused occlusions. The occlusions are when part of the vehicle is blocked by another vehicle. When shadows were present, the error could increase in some cases.

Based on the results of the feasibility study, the recommendations are to further improve the overall vision-based method such that it can cover a larger array of vehicles and overcome the challenges revealed in the feasibility study. The feasibility study was essential to remedy gaps in the existing literature and report the results of real-time processing. The lessons learned from the project developed the proposed framework discussed in Chapter 4. The proposed framework diverts away from typical vision-based algorithms for height estimations of objects and uses a more practical approach in detecting OHV. The proposed framework is not restricted to the shape and size of vehicle and the number of laneways in a traffic scene. Day and night capabilities does not impose on the performance of the framework.

Sample	Detected Vehicle Height (mm)	Actual Height of Vehicle (mm)	Accuracy	Truck Height Error (cm)	Raw Err (8cm correction)
1	3248	3365	96.52%	12	4
2	1998	2118	94.33%	12	4
3	2036	2223	91.59%	19	11
4	2109	2146	98.28%	4	-4
5	2109	2218	95.09%	11	3
6*	2894	2494	116.04%	-40	-48
7	2168	2660	81.50%	49	41
8	2468	2577	95.77%	11	3
9	3468	3514	98.69%	5	-3
10	1987	2099	94.66%	11	3
11	2106	2196	95.90%	9	1
12*	0	3368	0.00%	-	-
13	3190	3196	99.81%	1	-7
14	2155	2162	99.68%	1	-7
15	2561	2586	99.03%	3	-6
16	2234	2277	98.11%	4	-4
17	2036	2157	94.39%	12	4
18*	3056	2566	119.10%	-49	-57
19	2231	2255	98.94%	2	-6
20	2103	2313	90.92%	21	13
21	2657	2718	97.76%	6	-2
22	2425	2633	92.10%	21	13
23	4439	4602	96.46%	16	8
24	3136	3383	92.70%	25	17
25	3364	3622	92.88%	26	18
26	2739	2835	96.61%	10	2
27	3891	3944	98.66%	5	-3
28	3367	3447	97.68%	8	0
29*	0	1552	0.00%	-	-

Sample	Detected Vehicle Height (mm)	Actual Height of Vehicle (mm)	Accuracy	Truck Height Error (cm)	Raw Err (8cm correction)
30	2511	2683	93.59%	17	9
31	3069	3162	97.06%	9	1
32	2433	2567	94.78%	13	5
33	3891	3934	98.91%	4	-4
34	987	1256	78.58%	27	19
35	2469	2507	98.48%	4	-4
36	3967	4036	98.29%	7	-1
37	3109	3208	96.91%	10	2
38	2598	2664	97.52%	7	-1
39	2989	3057	97.78%	7	-1
40	2101	2133	98.50%	3	-5
41	2004	2068	96.91%	6	-2
42	3018	3023	99.83%	1	-8
43*	0	3846	0.00%	-	-
44	3093	3109	99.49%	2	-6
45	2136	2334	91.52%	20	12
46	3264	3327	98.11%	6	-2
47	3015	3111	96.91%	10	2
48	2368	2379	99.54%	1	-7
49*	0	2008	0.00%	-	-
50	2101	2116	99.29%	2	-7
51	3169	3201	99.00%	3	-5
52	2366	2446	96.73%	8	0
53	2169	2267	95.68%	10	2
54	2064	2097	98.43%	3	-5
55	2563	2630	97.45%	7	-1
56	3366	3460	97.28%	9	1
57	2709	2804	96.61%	10	2
58*	3026	2863	105.69%	-16	-24
59	3043	3056	99.57%	1	-7

Sample	Detected Vehicle Height (mm)	Actual Height of Vehicle (mm)	Accuracy	Truck Height Error (cm)	Raw Err (8cm correction)
60	3198	3328	96.09%	13	5
61	3167	3680	86.06%	51	43
62	3066	3067	99.97%	0	-8
63	2587	2601	99.46%	1	-7
64	3105	3122	99.46%	2	-6
65	2568	2684	95.68%	12	4

Table 14 Results of the data sample of the modified method.

Systems Specification
Multi-lane roadway
Two-way traffic
Day and night-time conditions
Accuracy ±5cm
Recall 100%
Precision 95%
False Detection < 2%
Warning Accuracy > 98%

Table 15 Systems specification with the major contributions of the system.

Systems Requirement	Minimum	Recommended
OS	Windows 7	Windows 10
CPU	Dual Core 2.4 GHz	Dual Core 2.8 GHz
VGA	GeForce 700 series better 256 MB	GeForce GTX 10 series 512 MB
RAM	2GB	4GB
HDD	1TB	4TB

Table 16 Systems requirement with minimum and recommended configurations.

The camera has built-in infrared capabilities and was tested during day and night-time conditions. During night-time conditions, the camera was better able to detect vehicles as less noise was visible (trees) and the object was more illuminated within the camera view resulting in better extraction of vehicles. During day-time conditions, the challenge was noise such as windy camera movements and swaying trees. Further work will need to accommodate these variables to increase robustness and stability of the camera.

Lessons learnt from this study reveal that we must design a system that is simpler. OH vehicle detection using computer vision does not require complicated vehicle segmentation or extractions to estimate the height of vehicles; there must be an alternative solution. The current method was too prone to 'mistakes' and detecting incorrect top and bottom boundaries of box-shaped vehicles. There are various shape and sizes of OH vehicles on today's roadway and we cannot restrict the problem to only a specific type of lorry. An improved algorithm must be able to accommodate for car carriers, flat trucks and/ or dump trucks to name a few. One way of tackling this problem is to look at the commonalities of OH vehicles and determine its uniqueness relative to the road. Table 15 shows the systems specification based on the lesson learnt. Based on the observations during the study, all OH vehicles exceed a certain height relative to the road plane. As part of the development, we can design a system to examine only the portion of vehicle over a certain height limit therefore minimising computational speed and processing. Further improvements are to accommodate for multi-lane streets and detection of vehicles on two-way streets. This can

be integrated as part of the systems requirement as per Table 16. The next chapter will discuss the proposed solution and how the system specifications fit in with the design of the system.

Chapter 4

Proposed Framework

4.1 Introduction

In this chapter, the proposed framework for the vision-based OHVD approach is introduced. A description of each of the key processes are defined in the trigger-based approach. Specifically, the selection process is discussed for selecting the ideal camera optics, placement and orientation for OHVD. The chapter concludes with the experimental set-up and validation, which explains how the framework set out here will be put to use.

The OHVD process is schematised in Figure 38. The basic concept of the OH detection method uses computer vision techniques to detect vehicles taller than a specific height relative to the roadway. The method uses a set of triggers to validate whether the vehicle is in fact, an OH vehicle. The detection is done by first, segmenting the video into image frames, which are then used as inputs for the OH detection process. If an OH vehicle is detected, recording by cameras and accelerometers is triggered; simultaneously, a message is displayed on the display unit, averting the driver to the presence of the relevant structure. The driver warning process may continue down one of two paths: 1) if no impact is detected (assumed to be due to the driver exiting the roadway), video data is discarded and accelerometers are deactivated; 2) if impact is detected, the camera captures the vehicle's number plate, the accelerometer records the impact data. The collision report, consisting of video footage plus number plate and accelerometer data, is sent on to the relevant authorities. This process is now explained in more detail.

Video is first converted into image frames. MATLAB's VideoReader is used to read video files. The video is then converted into individual image frames. After each frame is converted, it is passed to an image blur metric (Do, 2009). If the frame is identified as blurry, the code discards the frame and uses the succeeding frame. The blur metric works by passing images through several filters and assigning a 'blur annoyance' rating estimated using neighbouring pixels. If variation is high, the initial image is considered sharp. If the variation is moderate or low, the initial image is considered blurry. The blur perception is calculated using the sum of the coefficients and selected using the vertical and horizontal

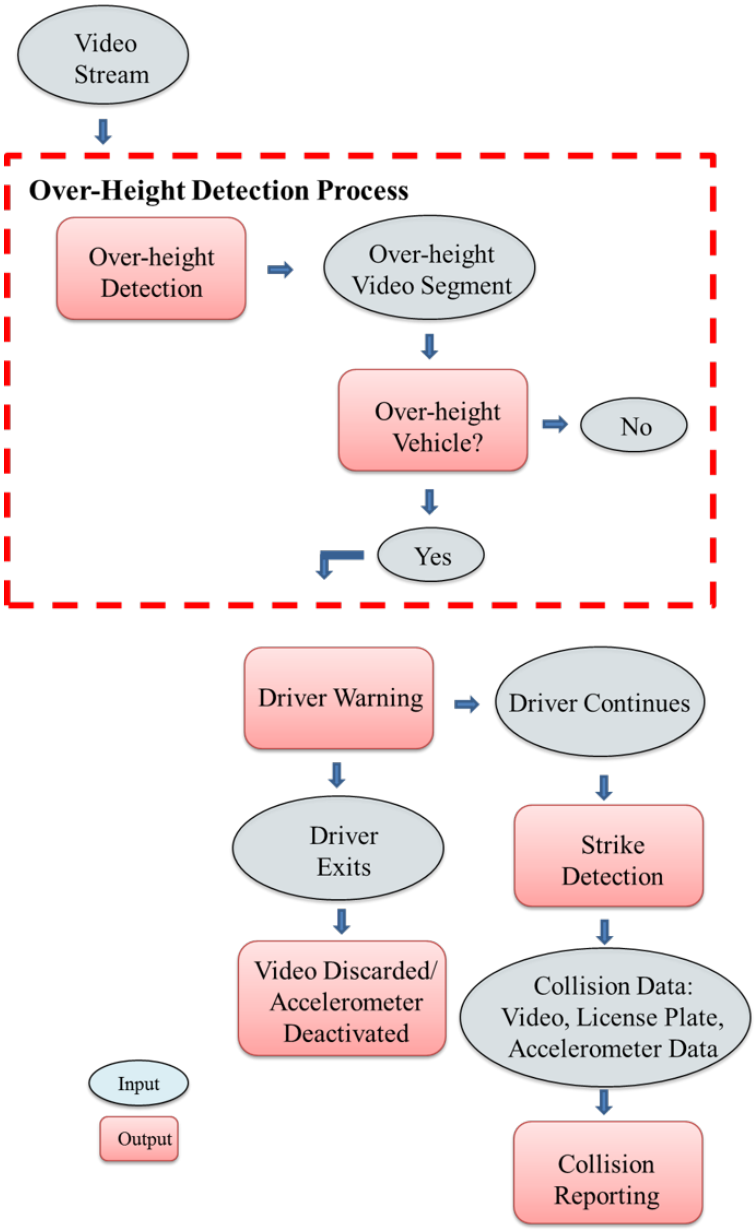


Fig. 38 Proposed framework for the trigger-based vision method for over-height vehicle detection, assuming pre-calibration of camera prior to video processing.

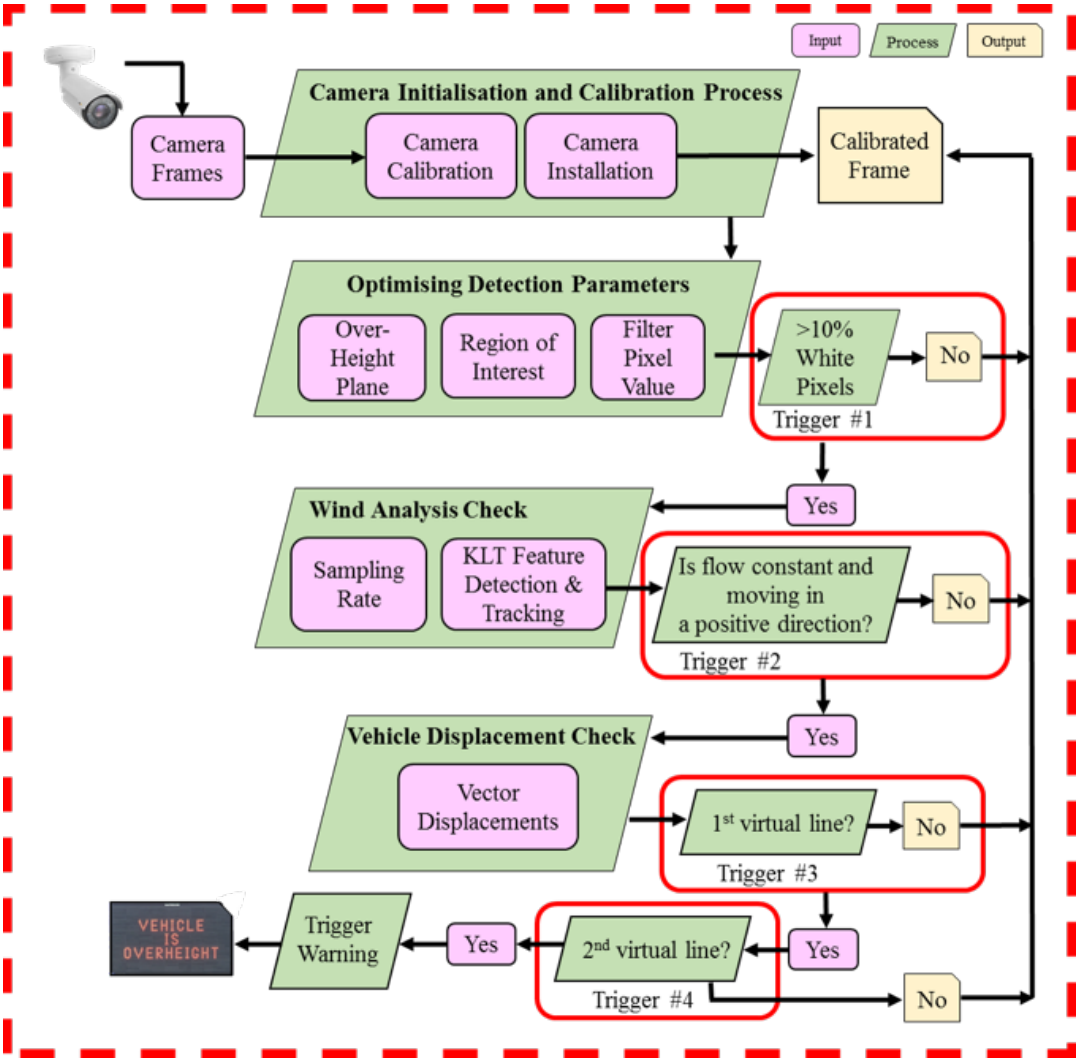
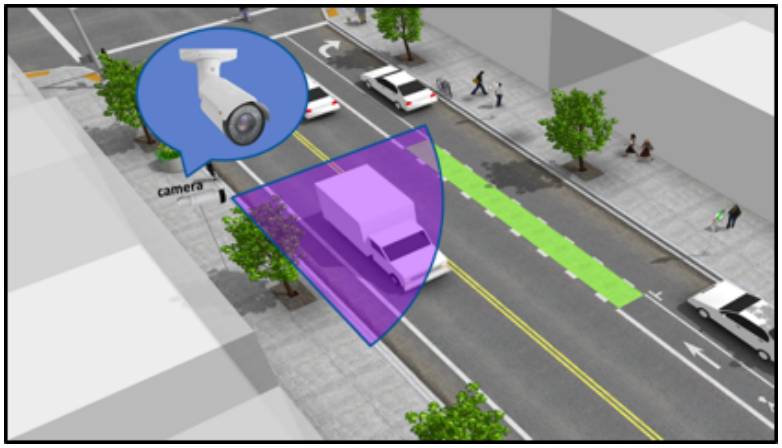


Fig. 39 A novel trigger-based approach for over-height vehicle detection and warning drivers of imminent low structures.

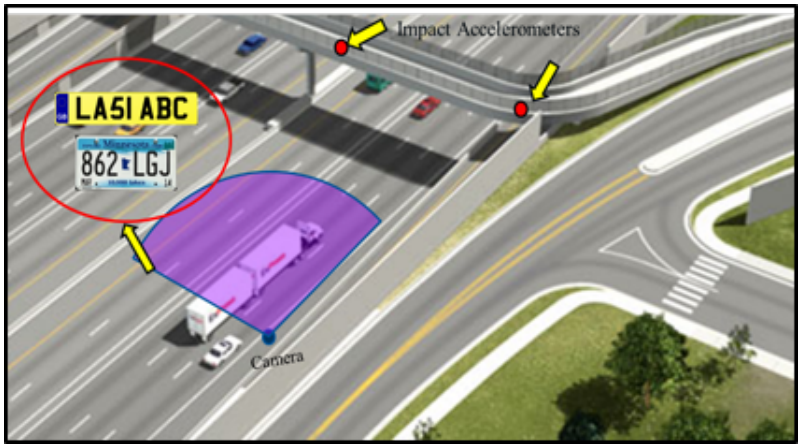
blur value, resulting in a binary solution for the sharp (1) and blurry (0) quality images. An OH vehicle is typically in the scene for 2 seconds for 30-50 mph roadways based on collected data. If the camera is set at 30 fps then ~60 frames are to be processed. For an alert to be triggered, consecutive frames of OH vehicles must be flagged. The system's persistence is based on a set of positive flagged frames to make a decision.

4.2 Over-height vehicle detection process

In Figure 39, inputs, processes and outputs within the detection and warning process are coloured purple, green and yellow, respectively. Four sub processes occur before any warning to drivers. These subprocesses are next described in more detail.



(a)



(b)

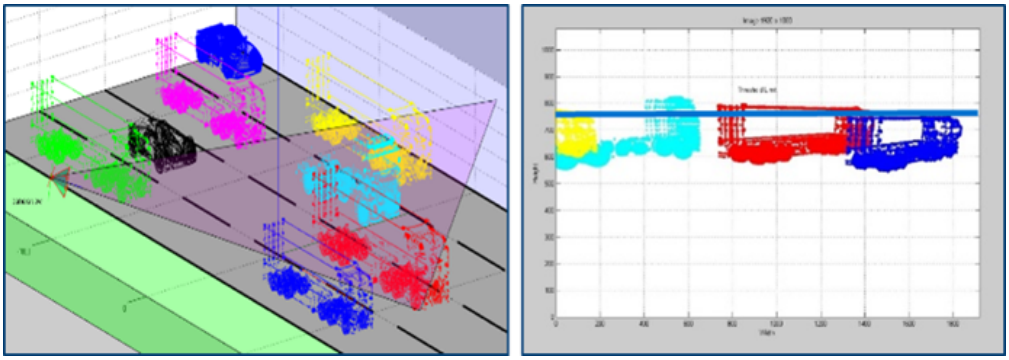
Fig. 40 Over-height detection camera in (a) urban setting and (b) highway setting.

4.3 Camera initialisation and calibration process

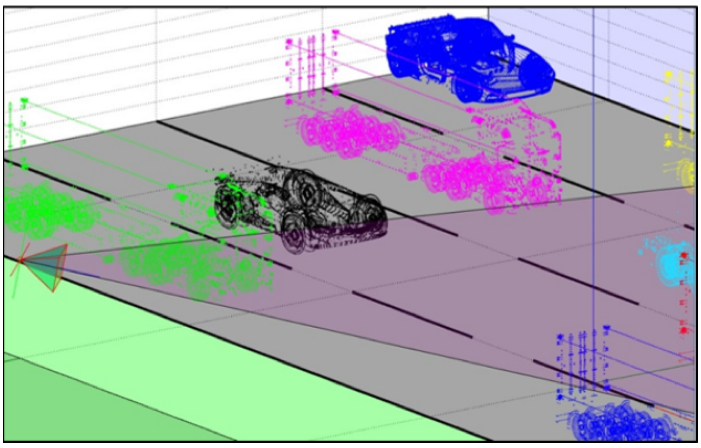
The calibration process involves the scene configuration and setup of the camera. In this section, a description of how to optimise the optics, placement and orientation of the camera for OHVD are presented. Figure 40 shows two scene configurations for OHVD in an urban and highway setting.

The camera is installed on the side of the roadway, approximately 1.5 m from the edge of pavement. The camera installation procedure accommodates multiple laneways despite the direction of traffic (further explained in Chapter 5). Impact accelerometers and number plate recognition are features of the system but are adequately treated in existing research and will receive no further mention here.

The camera calibration process is based on a simple geometric principle: the over-height plane which appears as a line in the view of a camera mounted at the height of the bridge clearance. The process thus mimics an active laser sheet using a passive vision method. The OH plane is constructed using three points $[x_i, y_i, z_i]$ set at the height h, representing



(a)



(b)

Fig. 41 (a) Camera configuration for multi-lane roadway; (b) Camera view with over-height plane and; (c) Close-up view of camera and over-height plane in a traffic setting.

the bridge clearance. The process requires the intrinsic and extrinsic parameters of the camera for setting the OH plane, as detailed in Chapter 5. In Figure 41, a camera is modelled using MATLAB to demonstrate the camera configuration and OH plane for a multi-lane roadway. As vehicles pass through the camera's field of vision, the detection algorithm analyses motion above the OH plane.

4.3.1 Ideal camera optics

Naturally, camera optics must be considered carefully in a vision-based approach such as the current one. The camera must be robust to standard outdoor conditions, but also severe weather conditions (snow, wind, rain). An outdoor camera housing unit is typically sufficient to prevent rain drops on the camera lens. The camera must have sufficient resolution for image processing with a zoom option in the event the camera is located far from the roadway. Sufficient camera pixel resolution is needed for number plate recognition, scene recording and collision reporting.

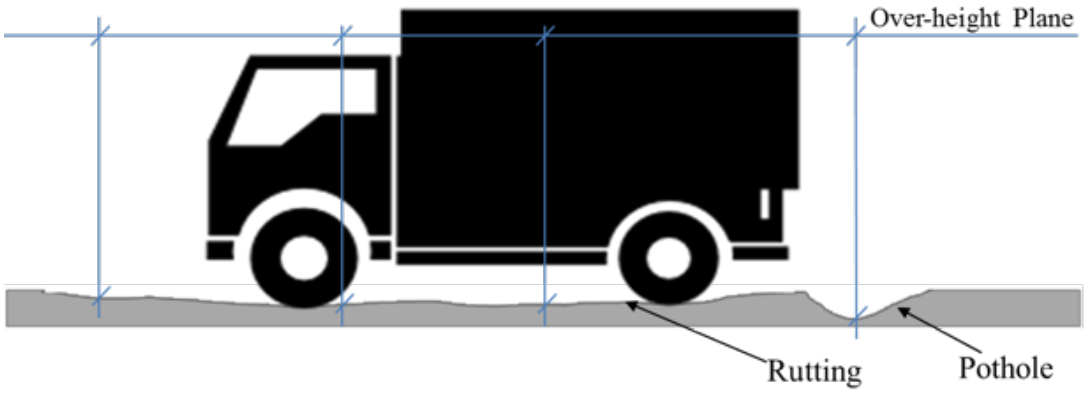


Fig. 42 Roadway with rutting and pothole. Such a scene selection should be avoided if possible, but where unavoidable, an average of the road heights is considered during calibration.

4.3.2 Camera placement

The camera is ideally set back 1.5 m from the roadway to avoid any potential damage from oncoming vehicles, and to allow for a wider field of view. Also, there is potential for an OH vehicle to occlude a subsequent OH vehicle if the camera is too close to the roadway. Ideally, the camera should be located on a straight roadway, where obstructions such as vegetation, parts of trees, and overhead cables are absent from the field of view. The camera should also be placed on a roadway without adjoining side streets intervening between the field of view and the relevant low structure, for the obvious reason that OH vehicles would then potentially be missed.

The roadway selected should be relatively defect-free i.e. with no potholes, rutting, cambers or slopes (up/down) to minimise errors during the calibration and detection process. In Figure 42, a scenario of the roadway with rutting and pothole are shown. If such defects are unavoidable, the camera calibration process considers the average of the heights relative to the road plane within the relevant segment in order to calibrate the OH plane. Accurate detection is thus still possible with a defective roadway, but is to be discouraged.

4.3.3 Camera orientation

Figure 43 shows two camera orientations: one at 90 degrees, and another between 30-45 degrees. The first scenario is for the simplest case, ideal for cases that do not require number plate, scene recording or collision recording. Due to the camera angle, visibility of the vehicle's number plate is not possible however, OH detection and warning are not compromised. In the second scenario, the setup allows for number plate recognition, scene recording and collision reporting when the camera is angled at between 30 and 45 degrees' parallel to the roadway. Although the optimal choice is the latter scenario, some asset owners may consider the former setup in less critical cases. For example, asset owners

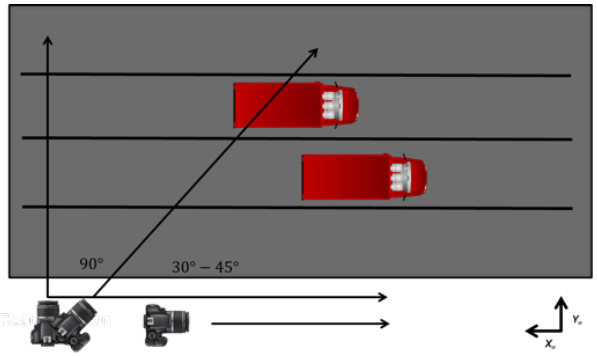


Fig. 43 Camera orientation at 90 degrees and 30-45 degrees respectively.

may have 10 critical sites; however, a limited budget and can only afford so many full installations. The former setup can act as a pre-installation setup to collect data of sites prior to committing to a full systems installation. The pre-installation setup can record the number of OHV instances within a specified time frame and that information can be used for decision-making. Technically, both set-ups can be performed as a pre-installation measure however, the former installation calibrates in less time and ideal for quick analyses (to test whether a site is worth the investment).

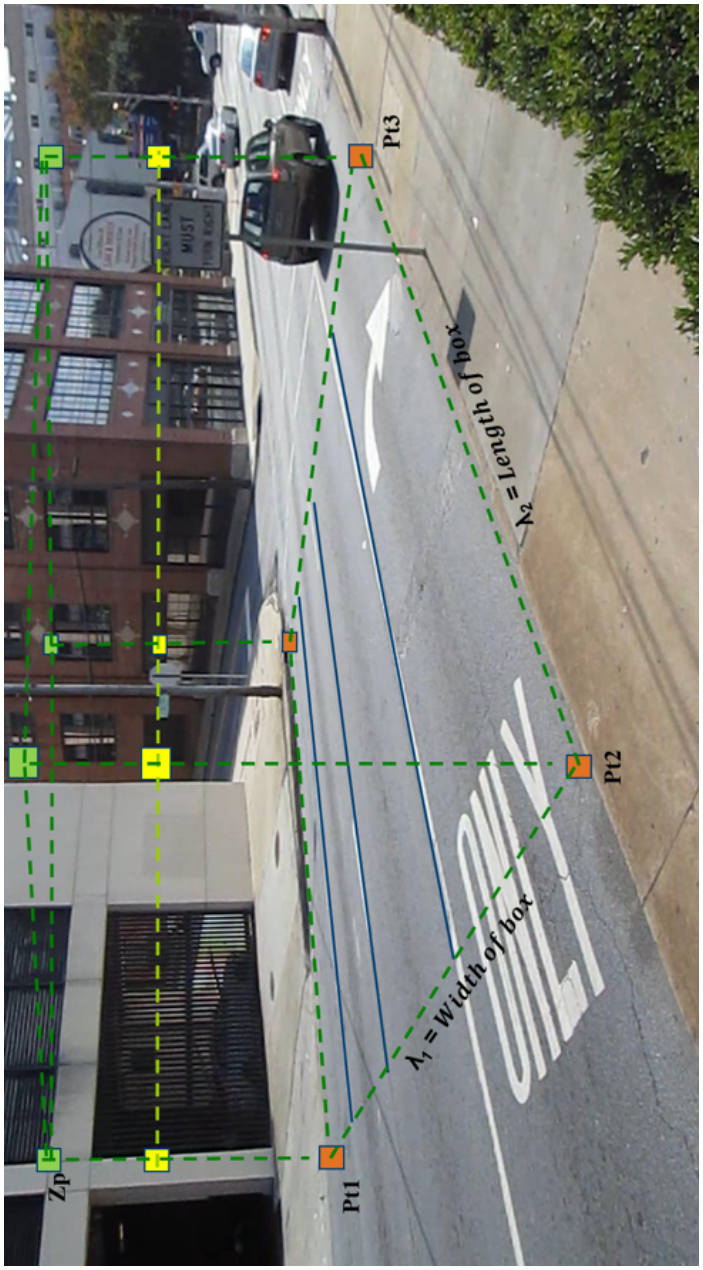


Fig. 44 Detection box to determine the camera resolution for number plate recognition.

Figure 44 shows a typical roadway with a detection box, plus the OH plane in yellow. The detection box dimensions are expressed as:

$$\begin{aligned} \lambda_1 &= \text{width of box (metres)} \\ \lambda_2 &= \text{length of box (metres)} \\ Pt_1 &= \text{coordinate position of point 1 } (x, y) \\ Pt_2 &= \text{coordinate position of point 2 } (x, y) \\ Pt_3 &= \text{coordinate position of point 3 } (x, y) \\ Z_p &= \text{height of over-height plane (metres)} \end{aligned}$$

$$(1-\lambda_1)Pt_1 + \lambda_1 Pt_2 + \begin{bmatrix} 0 \\ 0 \\ Z_p \end{bmatrix} \tag{7}$$

and

$$(1-\lambda_2)Pt_2 + \lambda_2 Pt_3 + \begin{bmatrix} 0 \\ 0 \\ Z_p \end{bmatrix} \tag{8}$$

which describes the width and length of the detection box. The height of the OH plane is the translation matrix expressed as $\begin{bmatrix} 0 \\ 0 \\ Z_p \end{bmatrix}$.

Anagnostopoulos et al. (2008) report that the camera for number plate recognition performs ideally between 0–30 degrees (parallel to the road plane) with a resolution of 728 x 1080 or higher and minimum plate resolution of 129 x 41 (by taking the averages of plate detection performance and minimum plate resolution). Based on the results gathered from the report, the minimum resolution is applied for the OH detection scenario for number plate recognition.

4.4 Optimising detection parameters

In this step, the image frame is initialised using optimised parameters. The parameters are designed to overcome noisy backgrounds that may lead to false positive detections. This includes moving vegetation, wind, rain, and any other illumination changes caused by lighting or weather phenomena. Detailed experimental setup and results pertaining to parametrisation can be found in Chapter 6 of this thesis.

Three parameters are relevant here: (1) the OH plane, (2) the region of interest (ROI) size and (3) filter pixel response value. Background subtraction is used to detect motion within the ROI. This motion is known as the foreground image displaying as white pixels in an

image. If the foreground image within the ROI exceeds 10% white pixels, the next stage of the algorithm, wind analysis check, is performed.

4.4.1 Wind analysis check

Wind analysis check is the first step in minimising the number of false positive detections. This step analyses motion within the ROI to discern true OH vehicles from noise. Noise is considered to be movement due to vegetation, illumination and so on.

Outdoor traffic cameras are prone to movement, which must be taken into account in the detection process. All motion captured within the ROI is analysed, being treated as a binary linear classification problem as follows:

$$Trigger\ Warning\ System = \begin{cases} 1, overheight\ vehicle \\ 0, noise \end{cases} \tag{9}$$

The algorithm is initialised when the flagged frame is passed to the wind analysis trigger. The algorithm uses a control variable to analyse the direction of movement using a feature descriptor. This control variable analyses the direction of flow movement and scores each feature-point. Sampling rate information is used as input to activate the feature-tracker detection algorithm. Any motion passing through the ROI is detected and tracked. Each point is tracked over x consecutive frames and analysed with reference to its neighbours to determine whether the flow is constant and moving in a positive direction. If the flow is constant (monotonically increasing) in the horizontal pixel dimension, the vehicle trajectory process is triggered (see below). If the motion is oscillating, the instances are classified as noise, and the process starts over.

4.4.2 Vehicle displacement check

The vehicle displacement process is the final pre-processing step to minimise false positives, and uses OH vehicle behaviour to set 'activation' regions (similar to motion detection regions). Motion detection regions is highly used in video surveillance systems to trigger an alert when motion is detected (Chandana, 2011). If motion is detected within these regions, the assumption is an OH vehicle is present.

This step uses two activation regions to trigger an alert if motion is detected. The assumption is that OH vehicles travel in a specific direction (easterly direction at distance d from the camera position) and secondly, at a constant speed as shown in Figure 45. The assumptions are used to set the activation regions within the frame, one region located at the left side of the image where the vehicle enters, and one region at the right side of the image where the vehicle leaves. If the vehicle crosses these two regions, motion activity is assumed to be OH- rather than noise-related.

(a) (b)

Fig. 45 (a) Shows the vehicle entering from the left side of the image and (b) exiting at the right side of the image.

4.5 Experimental setup and validation

The experimental setup tests separate components of the system to determine individual performance of each of the four sub processes described above. The validation stage then tests the integrated system.

The experiments are designed to: (1) calibrate the camera, (2) determine the optimised algorithm parameters for OH detection, (3) minimise the number of false positive detections endured by the system and (4) validate the system using the optimised parameters. The information is used as to determine whether the motion detected is due to an OH vehicle, and not to random noise. The initial experiment tests the system under ideal weather conditions (sunny, calm). The second experiment is designed to test the system under more rigorous weather conditions (rain, wind). The experiments aim to develop and test a vision-based approach to yield an accuracy of 5 cm or less.

The algorithm should deliver reliable and accurate real-time classification, as well as being robust in the face of frequent noise to provide accurate warnings to drivers. The designs of the experiments are categorised into four phases:

Phase (1) – Camera initialisation and calibration process (Chapter 5)

This camera initialisation and calibration process experiment is designed to calibrate and test the system to determine the accuracy of the OH plane in the image plane. The experiment explores calibration techniques for setting the OH plane. The experiment tests the proposed approach using various camera angles, road speeds and road geometry to determine optimal camera setup. Datasets from the US and UK are evaluated. Varying traffic regulations from the two case study locations help determine the optimal camera approach while considering roadway traffic, road speeds and camera angles. This ideal case will set the precedence for the following experiments. Further details of this experiment are given in Chapter 5.

Phase (2) – Determining optimised detection parameters (Chapter 6)

The experiment involves an optimisation procedure, which determines the ideal parameters to accommodate for variable illumination and weather conditions. The system is tested under sunny, cloudy and rainy weather conditions, under 25 mph wind activity. Once optimal parameter settings have been determined, further image processing is initialised. Experiment 2 is described in Chapter 6.

Phase (3) – Wind analysis and vehicle displacement checks (Chapter 7)

Experiments 3 and 4 are designed to minimise the number of false positive detections endured by the system. These final triggers exploit the behaviours of OH vehicles and use their characteristics to track the vehicles' displacement in the scene. Additional tests such as ideal frame and sampling rates are determined in this experiment. The frame rate is tested under various rates: 5, 10, 15, 20, 25 and 30 frames per second. The experiment determines the appropriate number of frames as well as the minimum effective frame rate sufficient for OH detection. Experiments 3 and 4 are described in Chapters 7.

Phase (4) – Validation of over-height detection process (Chapter 8)

The system is lastly tested as a package, with the four individual experiments combined. Validation results are compared with each of the individual experiments to determine the areas yielding the highest errors. This allows for further optimisation in areas that require adjustments. The validation approach is detailed in Chapter 8.

The experimental setup and validation aim to provide an optimal approach to the positioning of the camera, calibrating of the camera and determining the optimised parameters of the camera for OH vehicle detection. The lessons learnt in Chapter 3 helped develop Phase 3 which provides a thorough systematic set of triggers to minimise the number of false positive detections on the system. The complimentary checks, wind analysis and vehicle displacement are a set of addition parameters to increase the robustness of the overall detection process. Phase 4 intends to measure how the overall system performs in real-time.

Chapter 5

Camera Initialisation and Calibration

5.1 Introduction

Drivers of OHV often ignore warning signs and strike low-clearance bridges despite the presence of preventative measures. Managing bridge (and tunnel) strikes requires attention in three domains: prevention (discouraging strikes in the first place); detection (accurately recording strikes that do occur); and reporting (efficiently communicating OHVS details to the relevant authorities). The latter two aspects of OHVS management are effectively managed by current systems. Many OHVS technology currently on the market is preventative in nature. Very few systems are designed to mitigate OHVS impact (as opposed to preventing it entirely), as asset owners are interested in protecting the structure and limiting any risk of structural instability.

OHVS prevention systems can be categorised into passive, sacrificial, and active types. Asset owners tend to favour quick, cheap, and accessible passive methods such as signage, bridge markings, and flashing beacons to warn drivers. These are readily available, easily installed, and minimise additional infrastructure installation. Where strikes have persisted, practitioners may incorporate sacrificial or active systems. However, the biggest issues for asset owners are affordability and reliability, without compromising the accuracy and performance of such systems. Many systems are on the market; none cover all three aspects of OHVS management affordably. The chapter intends to answer:

> **Aim 1:** Determine the calibration process, scene configuration and setup of the camera including optics, placement and orientation.

The chapter proposes a viable solution for OHVD, specifically addressing the prevention problem. The chapter is organised as follows: Section 5.2 describes an overview of existing and related methods, followed by a general framework of OHVS management. Section 5.3 introduces the new proposed approach detailing the geometry, camera installation procedure, and detection algorithm. An evaluation of the system is presented in Section 5.4, including experimental results, discussion and 5.5 concluding remarks.

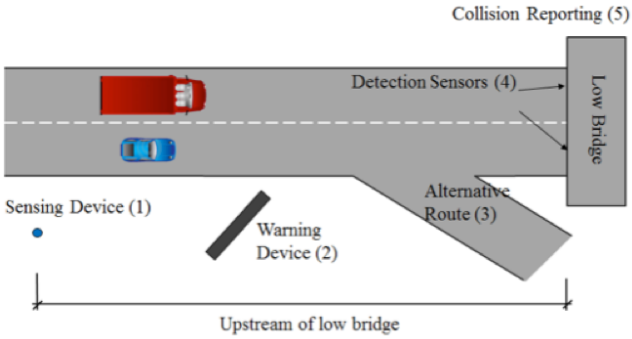


Fig. 46 Schematic layout of early warning detection system for over-height vehicles.

5.2 Background

Vehicle heights are continually increasing, and bridge structures built to standards that are decades out-of-date and often inadequate today, making OHVS an ongoing nuisance for asset owners.

Figure 46 depicts the *over-height vehicle detection* (OHVD) and warning system by Nguyen, Brilakis & Vela (2017). The system consists of the following components: sensing technology (1), warning device (2), alternative route to exit (3), detection sensors (4), and collision reporting (5). Components (1), (2), and (3) cover the prevention aspect of OHVS management using a sensor to detect the OHV and a warning device to warn the OH driver. Sufficiently low latency is required between data processing and warning issuance to provide the OH driver with sufficient time to react, brake and exit. In ideal situations, an alternative route is provided for safe exit. Components (4) and (5) cover the detection and reporting aspects of the system. Detection sensors are mounted on the bridge structure to record any vibration or impact frequencies caused by strikes, and real-time collision reporting is used to notify authorities of strikes. The system thus represents a holistic early warning and detection system for OHV.

Existing vision-based methods have been discussed in Chapter 2; there, it is made clear that preventative methods are an active area of research, with the perennial goals of high performance and low cost. This section reviews this research in more detail, concentrating on preventative methods based on imaging or electromagnetic waves. Imaging- or vision-based sensing is divisible into two categories according to sensor modality. The first involves sensors with active illuminators or emission of electromagnetic waves; lasers and radar are prominent examples. The second involves sensors that passively measure the ambient electromagnetic energy, the standard video camera being the main example. The review of passive prevention methods will be further split into active sensing and passive sensing strategies.

Active sensing methods consist of optoelectronic single- or dual-eye infrared, visible beam, radar or laser beam detection systems, all of which detect OHV when the laser or

light beam is interrupted (Sinfield, 2010). Such sensing technology methods, however, provide little incentive for asset owners, since the infrastructure requirements are financially prohibitive. Outdoor infrastructure entails the installation of new permanent poles, usually a receiver and transmitter for laser-based cases.

Passive sensing uses vision-based methods similar to those currently used for vehicle detection, vehicle classification, and number plate recognition (Anagnostopoulos et al., 2006). Pertinent to these tasks, scene change detection (background subtraction), vehicle tracking and motion detection, all of which are essential for OHVD, have been the subject of much research (Coifman et al., 1998, Jazayeri et al., 2011). Researchers have studied alternative approaches using vision-based methods to extract vehicle height measurements, but no vision-based system is available on the market. Research specific to OHVD has been somewhat limited, but provides a solid starting point for further development.

Other computer vision methods for OHVD are found in Kanhere and Birchfield (2008), Shao et al. (2010), Criminisi et al. (2000), and Sturm and Maybank (1999). All these authors use vanishing lines and reference objects to extract height measurements of vehicles and objects. These passive methods all rely on the same underlying concept: given a known ground plane and an upper and lower limit, one can recover the height of objects in the scene. However, relying on geometric shapes and structures to recover usable information in complex scene may not be the easiest approach. OH in essence means the object (or subject) is in excess of height. Therefore, the problem can be approached by minimising the set of confounding factors by understanding the cameras position and vehicle behaviours such as motion. The next section discusses the methodology of the proposed solution.

5.3 Proposed approach

Existing *early warning detection systems* (EWDS) are the most accurate warning systems, yet are not cost-effective due to their significant physical infrastructure requirements. Cost considerations drastically limit their adoption and suitability. New EWDS are needed that can bring the cost down by at least one order of magnitude to make them attractive to infrastructure owners. Therefore, this section presents a novel framework for OHVD using perspective projection, inspired by the laser beam method. The objective is to replace the laser beam method consisting of a transmitter, receiver, and loop detectors with a single camera mounted upstream of a low bridge.

The proposed method (Nguyen, Brilakis and Vela, 2017) is based on the following geometric principle: when a camera is properly mounted at the height of the bridge clearance relative to the local roadway, the OH plane will appear as a line in the camera image. The method is suitable for various shapes and sizes of vehicles, numbers of laneways, and illumination conditions (day and night). Camera placement is crucial; this step minimises any potential captures of noisy motion that may contribute to triggering false positive alarms. The roadway around the camera location should be free of potholes (to minimise

height variations), vegetation, tree branches, and overhead cables. In line with the modelling of perspective projection, if objects are below the set camera height, they will not be detected within the ROI despite distance from the camera (this includes building occupant motions from across the roadway). However, if building occupants are on the second floor and captured within the ROI, the practitioners should find an alternative location to minimise noise in the data. If an alternative location is not possible, the threshold needs adjustment to be adjusted to account for the noise (further explained Section 6.4 using optimised parameters).

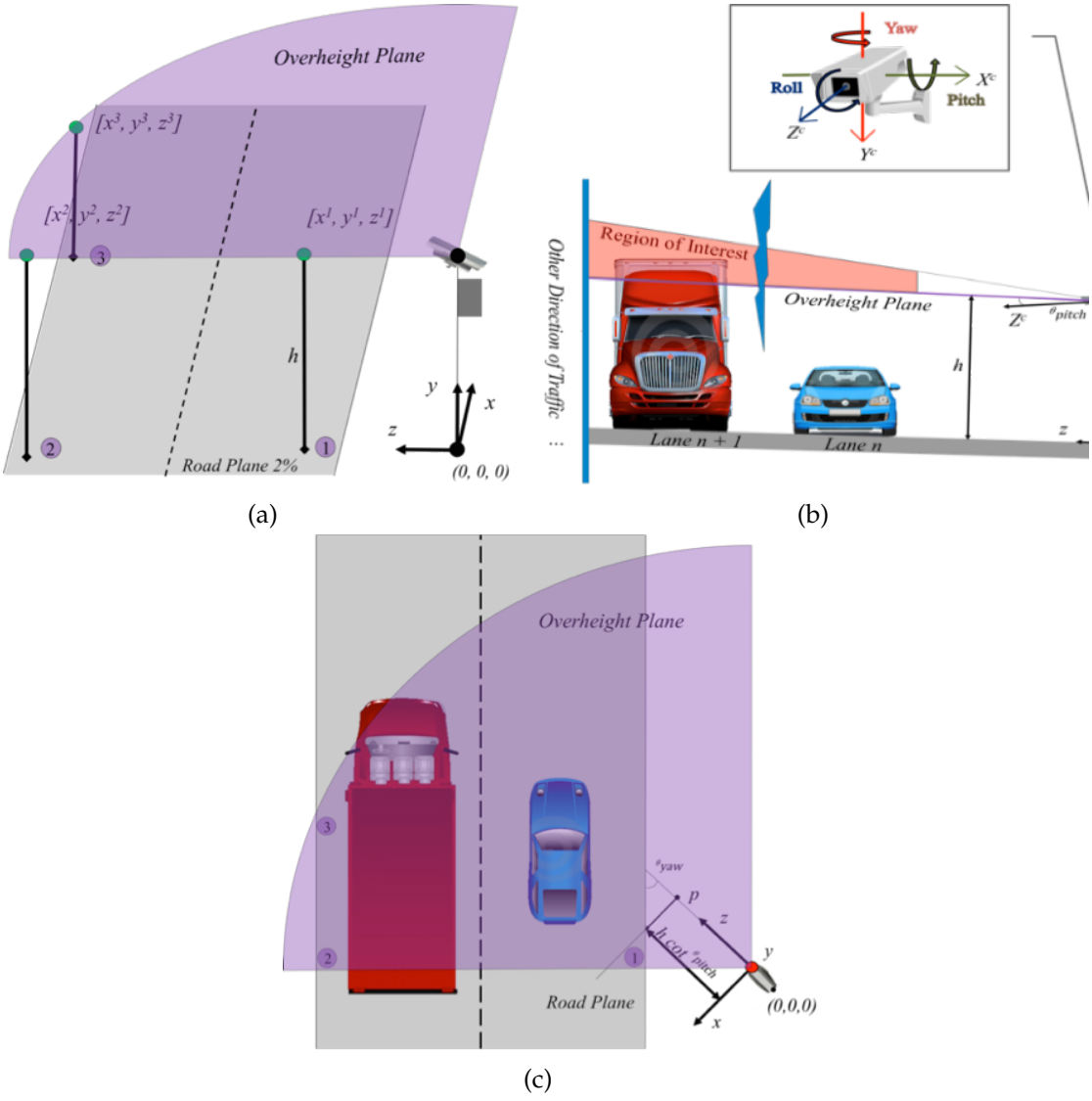
The primary innovations described here are the specialised camera placement relative to the roadway, and the associated setup procedure that minimises installation effort. All components of the system described are intended to minimise inspection, maintenance and repair costs. If the proposed solution can achieve the accuracy of laser-based systems while maintaining the low cost of passive vision-based systems, then a pairing with complimentary detection and reporting methods will provide a holistic solution to OHVS.

5.3.1 Camera geometry and detection policy

The method models an active laser sheet using passive vision methods. Figure 47 displays a cropped version of the infinite OH plane which is offset from road plane by bridge clearance height h. The camera coordinate system is X_c, Y_c, Z_c, and the world coordinate system is X_w, Y_w, Z_w.

The camera rotation is given as $\theta_{yaw}, \theta_{pitch}, \theta_{roll}$. The OH plane is obtained by offsetting the road plane by the height h, and the camera is placed such that the optical centre lies on the plane. The light rays of object points located on the OH plane will project to a line in the image. The plane divides the world into two regions, those above and those below. Correspondingly, the line in the image divides it into object points below or above itself. The method assumes that the lanes are approximately planar across the road width in each direction, that trucks keep left except to pass (UK), and that camera lens distortions are rectified through camera calibration.

Figure 47b depicts a side view of the OH scenario with an OH ROI (indicated in red). The pitch of the camera corresponds to a downwards tilt, ($\theta_{pitch} \geq 0$), to minimise reflection of sunlight onto the lens. This volume (ROI) projects onto the image as a band. Any OH vehicles passing through the scene will exceed the height of the line in the image view and project into the band, triggering an OH detection. In this sense, the proposed geometric setup resembles that of an active laser sheet. Figure 47c displays a birds-eye view of the camera setup. The optical axis of the camera Z_c intersects with the road plane along the y-axis at $p = (0, h.cot.\theta_{pitch}, 0)$. All figures use the right-handed system, such that x and X_c are into the page in the side view, while y is coming out of the page in the birds-eye view of Figure 47c, represented by the red dot. Next, the camera installation procedure is explained more in detail.



(a) (b)

(c)

Fig. 47 (a) Side view of camera orientation (b) Side view of over-height scenario (c) Birds-eye view of camera setup.

5.3.2 Camera installation procedure

This section summarises the mechanics of our method. There are two aspects to the calibration process, involving the intrinsic and extrinsic parameters of the camera. The intrinsic parameters are constants that hold irrespective of the placement of the camera, while the extrinsic parameters are tied to the placement of the camera in the world. Installation requires the extrinsic parameters to be determined according to the local roadway and the desired OH value h. However, intrinsic parameters should be established first.

The intrinsic parameters, being independent of placement, can be estimated anywhere. We propose that they be done away from the installation site, as the necessary calibration equipment may then be better controlled for accuracy. The standard method for intrinsic parameter calibration involves a calibration pattern. Photographing this calibration pattern from different positions and orientations enables the estimation of the intrinsic components of the camera such as focal length (f_x, f_y), camera centre (c_x, c_y) and radial distortion coefficients (k_1, k_2) (two coefficients are typically sufficient for compensation for radial lens distortion; see Heikkila and Silvén, 1997).

The extrinsic parameters represent the transformation from the 3D world coordinate system to the 3D camera coordinate system centred at the optical centre; together, the two sets of parameters (extrinsic and intrinsic) describe the transformation from 3D world points to 2D image points (Fathi and Brilakis, 2014). The camera installation and extrinsic calibration process configure the OH system in a controlled and replicable manner. The process relies on the facts that installation involves specification of two variables, camera height h_c and camera roll θ_{roll} and that a plane can be defined by any three non-collinear points lying on the plane.

5.3.3 Camera installation and extrinsic calibration process

The camera installation and extrinsic calibration process involves manipulating the projection of three predetermined OH plane points until the three points align on the same axis. Figure 48 represents a visual narrative of the installation process. The red arrows indicate the corrections needed. Consider Figure 48a, which depicts three non-collinear points $[x_i, y_i, z_i]$ set at the height of the bridge clearance h relative to the local roadway. The light rays that make up the plane project onto the image view as three $[x_i, y_i]$ points. When correctly installed, they will project onto a horizontal line in the image (the desired OH detection line); we refer to this simply as the 'OH line'. Initially, this will not be the case. The installation process provides a means to arrive at a horizontal OH line with height equal to that of the bridge clearance (here mimicked by a tall pole with a bright marker at the tip).

The camera is assumed to be installed at the height h above the road plane, and that the projection to the road plane is the road plane origin (0, 0, 0). First, the camera is placed on an existing utility pole (owned by the asset owner or permission granted by authorities) at an approximation of the desired height. Placing the camera on a pole limits

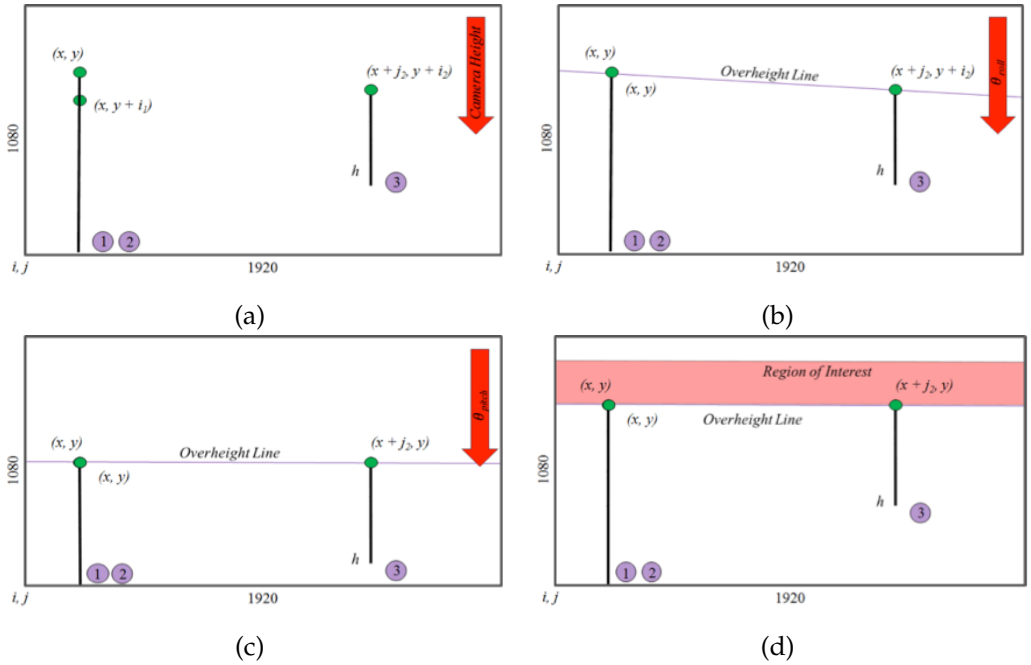


Fig. 48 Three points of the over-height plane are projected onto the image plane (a) The points are out of sync; corrections to camera height are required to align the over-height line in the image; (b) Over-height line is skewed; corrections to roll is required; (c) Adjustments of the pitch is required to optimally position the over-height line and region of interest; (d) Final position of camera and desired region of interest. Connecting the three image points (X_i, Y_i) forms the line for over-height detection.

the translational degrees of freedom to one. Then the following two rotations are set: θ_{yaw} is angled towards the road to capture number plates, and θ_{pitch} is tilted downwards to allow for optimal positioning of the ROI, and less illumination interference. By performing these two rotations, the user has met two of the three rotational conditions: θ_{yaw} and θ_{pitch}. Therefore, one degree of freedom θ_{roll} remains.

At this point, the user must perform two pole measurements. For the first point, the user should capture a measurement towards the left side of the image. The second pole location should be located behind the first, which is achieved by walking away from the camera along the line defined by the camera installation point and the first pole point (both projected to the road plane). The simplest way to do this is to face the camera, then walk backwards with pole in hand. If the camera is at the pole height, then both of these pole locations will have the pole tip marker project to the same point in the image. If not, then there will be an offset determined by the true height of the camera relative to the desired OH plane. If it is below the OH plane, then the first point will appear "above" the second point and the camera should be lowered; this situation is depicted in Figure 48b with the red arrow denoting the correction to be made. If it is above the OH plane, then the opposite will hold. The measurement and adjustment process should be repeated until the two pole tip markers project to the same point on the 2D image.

The camera will now be located at the proper height, but the OH line in the image will be skewed by a degree determined by the camera roll relative to the road plane. Hence, the θ_{roll} is modified. This step is not technically necessary, but it is recommended, as it simplifies OH detection computations. The user takes a third measurement which projects to the right-hand side of the image. The further to the right, the more sensitive the roll estimation process will be, and hence the more accurate. If the camera is at the correct roll, then this third point will lie on the same horizontal line as the first two points (their y-coordinate on the 2D image will be identical). Otherwise, the line defined by the projected image coordinate of the first two pole tip points with the third will have a slope. A positive slope requires clockwise roll adjustment, and a negative slope requires anti-clockwise roll adjustment. The scenario is depicted in Figure 48b. Some iteration may be necessary to arrive at the proper θ_{roll}, as depicted in Figure 48c. For each iteration, two points are needed, meaning that two pole tip measurements are required: one on the left side of the image and one on the right side, as depicted by marker locations (2) and (3) in Figure 48d, respectively.

The camera is now located at the proper height, with the necessary roll for a horizontal OH line. However, this line may be located too low in the image. A low placement means that the camera is capturing too large a proportion of OH region. While this is theoretically unproblematic, there are illumination factors to consider. Too much sky in the image leads to automatic exposure compensation that would darken the roadway in the image. Adjusting the camera pitch to minimise areas of sky thus should improve visual processing by minimizing confounding factors. At this point, the user can adjust θ_{pitch} so that the OH detection line creates a correct partition of the image, while still allowing for detection of OH vehicles within the determined region.

5.3.4 Detection procedure

The detection procedure uses video captured from the camera, continually converted into image frames, as its initial input data. Motion segmentation is used as the feature extraction method to detect and track moving objects within the ROI. The ROI is an area of pixels above the OH plane in the image, appropriately sized to minimise the risk of false detections. The detection algorithm calculates the motion differences within the ROI between the current image frame and background model by using vehicle motion when OH vehicles are present in the scene. Motion is detected by calculating the vector difference (optical flow) between the current image frame and background model, as shown in Figure 48.

OH feature points are automatically detected and tracked using the Kanade-Lucas-Tomasi (KLT) feature detection algorithm (Lucas & Kanade, 1981). Feature points in computer vision are specific structures in the image such as points, edges, corners or objects. There are many different ways to define feature points and to characterise them for image processing algorithms. Among these, the Kanade Lucas Tomasi (KLT) method is particularly suitable to this application because it allows for feature point tracking of vehicles within

image frames. The feature detection works ideally with static cameras as the rotation and pose estimation no longer poses a problem as the camera is not moving. The green circle represents the initial detected feature point detection i in the image, while the red cross represents the motion of that same detected point in the next consecutive image, $i++$. If no motion is detected, the circle and cross coincide. If movement is detected, a velocity displacement arrow (here in blue) is overlain showing direction of movement.

The camera setup allows OH vehicles to appear within the ROI; any moving objects traveling at a constant velocity in the direction of traffic are thus detected and tracked as shown in Figure 49. A motion threshold value is determined by comparing the pixel differences and adjusting for sensitivity to noise, including moving objects such as trees that may interfere with the detection procedure. Vehicle occlusions and shadows do not interfere with the detection process since the camera is situated at a height where occlusions and shadows are less frequent. For example, if the bridge clearance height is 6.0 m, then the ROI only detects vehicles over the height of 6.0 m. However, vehicular occlusions may occur when two or more OH vehicles are in the scene simultaneously; this occurrence will trigger a single undifferentiated warning to both drivers. Vehicle shadows are generally on the road plane out of range of the ROI, and therefore pose no problems to the detection procedure. The other set of uncontrolled environmental drawbacks consists in variable weather conditions (wind, rain cloud). The degree to which such environmental factors can interfere with the algorithm's performance has necessitated some extension to the present work, as detailed in the following chapters.

The detection procedure is suited to various shapes and sizes of OH vehicles. As long as some component of the vehicle exceeds the OH line in the 2D image, computing exact height measurements is unnecessary; rather, a binary decision (OH / non-OH) is sufficient for the desired functionality. The camera geometry and its associated installation procedure overcomes several deficiencies associated with existing detection methods. In particular, it eliminates the requirement of a vision-based ground plane measurement, which most of the other solutions require. Further, since visual processing focuses only on the portion that is OH on vehicles, the set of confounding factors is less than the current strategies, allowing for more efficient computation. Other methods return height estimates of all vehicles, therefore increasing computational time and processing efforts.

5.4 Experiment and results of plane calibration concept

This section provides details of the initial experiments designed to evaluate the height and detection accuracy of the system. The experiments were conducted on two collector roadways with 2 and 4 lanes of traffic in sunny, cloudy and rainy weather conditions. A Canon EOS M camera is used to capture 150 minutes of video data with 1920 x 1080 resolution at 30 fps. The processing unit is an Intel Core i7-4790. A camera is mounted on a fixed pole, with $\theta_{yaw} = 45°$ and $\theta_{pitch} = 10°$ set to capture number plates and to minimise

(a)

(b)

Fig. 49 (a) Camera system installed at Fairlop Underground Station, London (UK). The camera is installed at the same height as the bridge clearance; when calibrated, the camera defines the over-height plane and detects moving objects above the plane. The over-height/ 2D line in camera view and region of interest are displayed in yellow and red respectively. The region of interest is optimally positioned above the over-height to minimise processing requirements. (b) Zoomed-in view—moving over-height features are detected and tracked using the Kanade-Lucas-Tomasi algorithm.

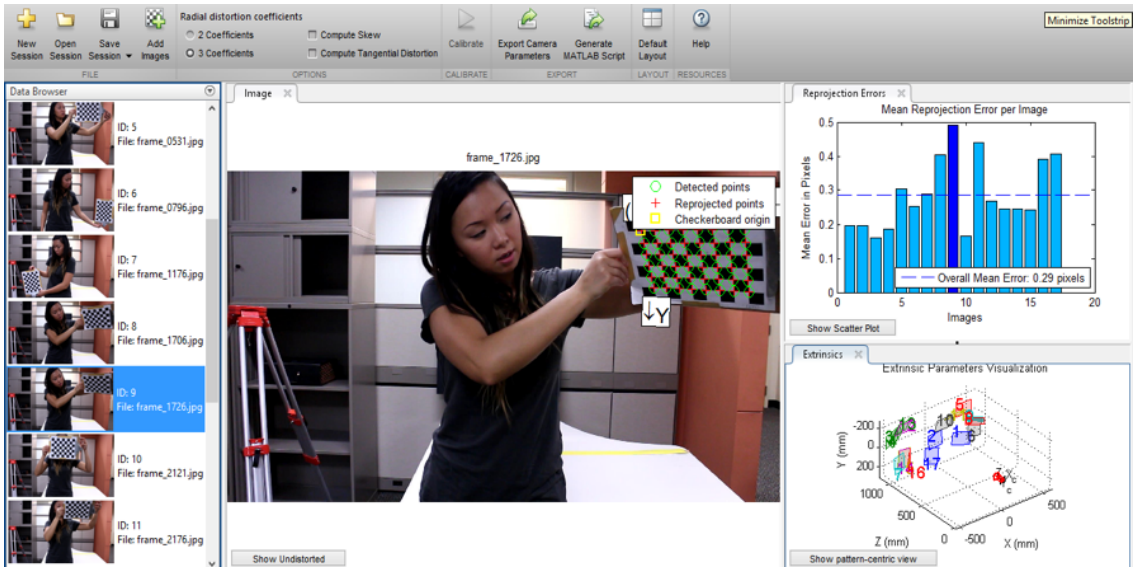


Fig. 50 Intrinsic camera parameters obtained with MATLAB's Computer Vision Toolbox (single camera calibration).

sun glare on the camera lens. The camera is installed 1 km upstream of the low clearance structure at a height of 5.0 m, to allow for: (1) detection of the OH vehicle, (2) issuance of warning message, and (3) sufficient time for the driver to react and to take the nearest exit. The camera is located such that no obstructions (excessive vegetation, trees, branches, overhead cables) were in the camera's field of view. The camera is set back 1.5 m from the roadway, to avoid potential damage from vehicles, as well as to allow for a greater field of view (inadequate distance of camera from roadway risks occlusion from vehicles). The roadway selected is relatively planar; no potholes or rutting were present, meaning errors during calibration and detection could be minimised.

An 8 x 6 mm calibration checkerboard pattern with 26mm squares are used as part of the intrinsic calibration process, as shown in Figure 50. MATLAB's single camera calibration process is used to find the intrinsic matrix ψ and two radial distortion coefficients k_1, k_2. These parameters were then used to undistort the images in order to find the correct X_i, Y_i points on the image plane.

Extrinsic calibration was performed using an extensible window washing pole set at the height of the bridge clearance h, with an attached prefabricated levelling bubble set plumb to the road plane. The OH plane was determined from the pole heights relative to the road plane; the error caused by the road gradient was assumed to be negligible and is absorbed through the calibration process. The road gradient under the UK Department of Transportation's road design specifications require a minimum of a 2% road slope (rise to run ratio) for sufficient water runoff to outlets such as catch basins, ditches, and culverts. This process considers the road gradient, whether or not the poles are parallel to the road surface's normal direction. For example, if the road grade is on a decline the camera will be tilted to the same degree, as the calibration process will align the OH plane with the grade

Fig. 51 (left) Total, total station method:) Leica TPS1200+; (right) Pole, pole method:) tennis ball attached to extensible pole with known height.

of the road. The following section compares two extrinsic calibration methods the: total station and pole, to determine its millimetre level of accuracy for setting the OH plane.

5.4.1 Accuracy comparison: Total station method vs. pole method

Setting the OH plane requires validation data to verify whether the height of the plane is accurately matched to the height of the low bridge. Validation data refers to the process of comparing real-world measured dimensions (measured on-site) with the measurement in the image in pixels. To perform this measurement, the accuracies of two methods were compared: the total station and the pole, as shown in Figure 51. The total station is the preferred method, with a distance and height accuracy of 1 in 1000; however, total stations are expensive and require specialised training to operate. The second method is the pole method. The pole method is an inexpensive alternative and since the height of the object is known, we can use the method as ground truth data. This allows for a quick and simple check. However, for OH detection purposes, it is unknown whether the pole method is sufficient. Therefore, a comparison of the two methods is needed. In this section, an experimental comparison of extrinsic parameter techniques is thus performed, in order to compare the height accuracy of each method.

In this comparison, we assume the height of the low bridge is assumed to be 2.40 m. In order to verify that the line in the camera view is accurately set to 2.40 m, the two aforementioned methods were used to measure this height. In the next two sections, the experiment and results are detailed.

5.4.2 Total station method: Accuracy and results

The experiment is performed on a 4-lane roadway in Atlanta, Georgia (USA). The camera height is set manually to 2.40 m, with 1920 x 1080 resolution. The purpose of this procedure

(a)

(b)

Fig. 52 (a) Visual representation of the points taken using the Tsai camera model and; (b) Projected image points with respect to the manually selected 2D image points.

	Corner 1	Corner 2	Corner 3	Corner 4
Distance	5689 mm	14222 mm	17034 mm	9482 mm

Table 17 Distance of pole from camera at each corner.

is to determine the mathematical relationship between the points in 3D and their projection onto a 2D image plane. The total station method will determine the accuracy of the camera height however, firstly the camera is calibrated. Tsai camera model calibration method was used, detailed earlier in this chapter. A total of 100 points were collected with a total station to compare 3D points X_w, Y_w, Z_w (real world locations) with the respective 2D image point locations x_i, y_i. Figure 52 shows the locations of those points taken from a traffic scene.

The 100 world points were used as inputs for the Tsai camera calibration model. The output of the calibration method showed a re-projection error of 4.49 pixels (Figure 53). The estimated height of the camera in the real world is 2.33 m; a difference of 0.073 m from the nominal measurement of 2.40 m.

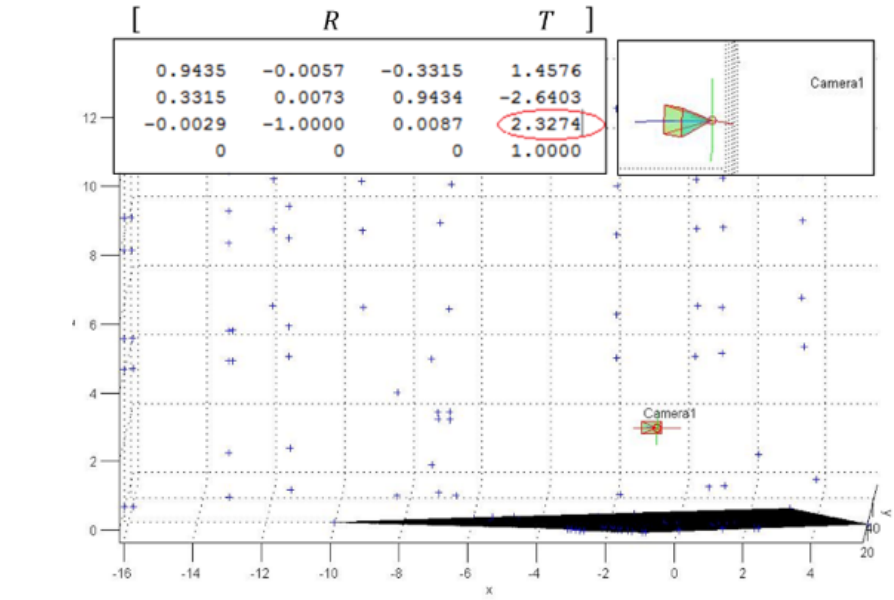
5.4.3 Pole method: Experiment and results

The pole method is performed on the same 4-lane roadway described above. The camera height is manually extended to 2.40 m. Table 17 shows the height accuracies between pole versus total station methods. The pole is constructed from an extensible window washing pole with a tennis ball attached to the top. The roadway is a typical crown-constructed roadway meaning its highest point is in the centre of the roadway. This is represented in red in Figure 54a. The pole is positioned at four corners of the roadway, forming a box as shown in Figure 54b.

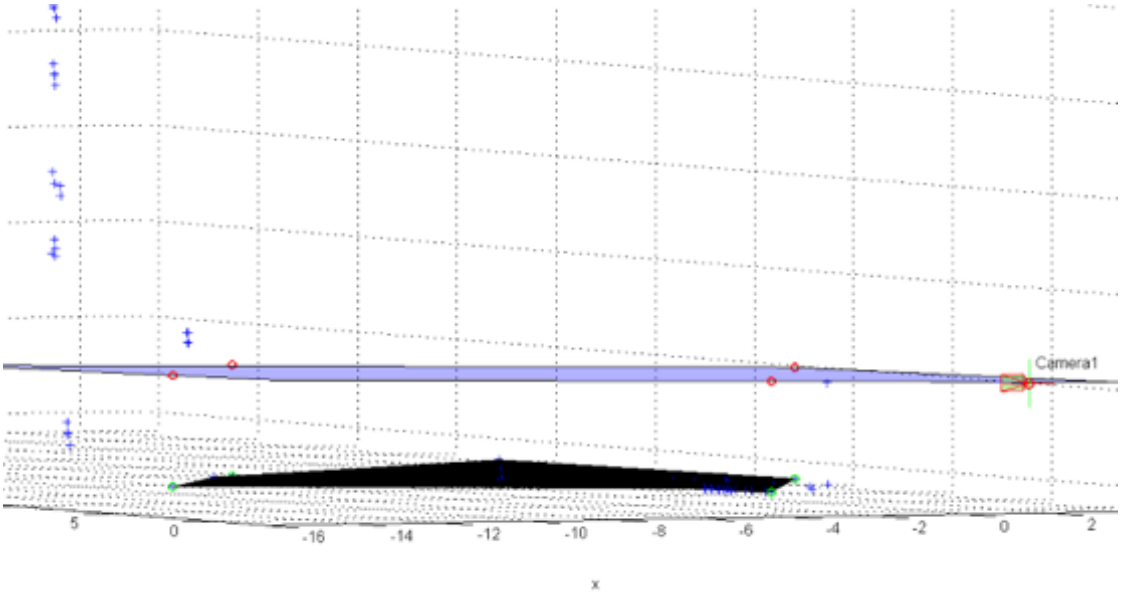
The demonstrator is instructed to walk along the rectangular box and stop every 0.5m to take a reading. Once stopped, the demonstrator is instructed to swing the pole in both the orthogonal and parallel directions to the camera to capture the pole vertically upright (this is done manually using a split bubble attached to the pole). The results show there is a ±12-pixel difference on the image (Figure 55) between the closest (point 1) and furthest (point 3) points, with their respective distances to the camera shown in Table 18.

The video taken of the pole at each of the four locations is decomposed into individual image frames before being undistorted. Each image frame is processed manually to find 2D positions (x, y pixel locations) of each of the pole locations (using the tennis ball in particular), and is then projected back onto the image plane as shown in Figure 56. The results determined the OH plane to correspond to pixel 569 on the y-axis. Table 15 shows the measured pixel locations of each of the points surveyed on the roadway.

A software installation prototype is created to aid users in performing the camera corrections needed in order to locate the three (X_i, Y_i) points in the image view. The prototype functions by retrieving and undistorting a single image taken when the poles are



(a)

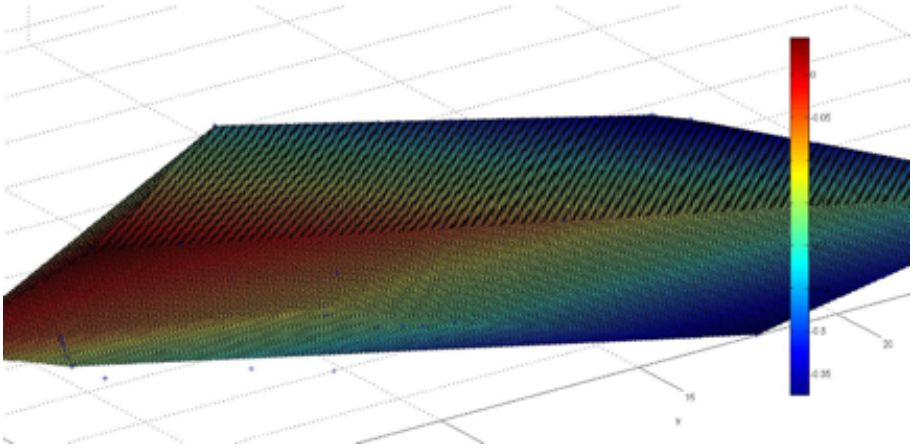


(b)

Fig. 53 (a) Tsai camera model calibration shows a translation height (T_z) estimation at 2.33 m, with the over-height plane estimated in the camera view; (b) Shows the camera projecting the over-height plane in 3D.

Location on roadway	Measured location	Pixel	Total points
middle of road (crown)	highest	575	23
edge of pavement (EOP)	lowest	563	15
	average	569	

Table 18 Measured pixel locations of each of the points on the roadway.



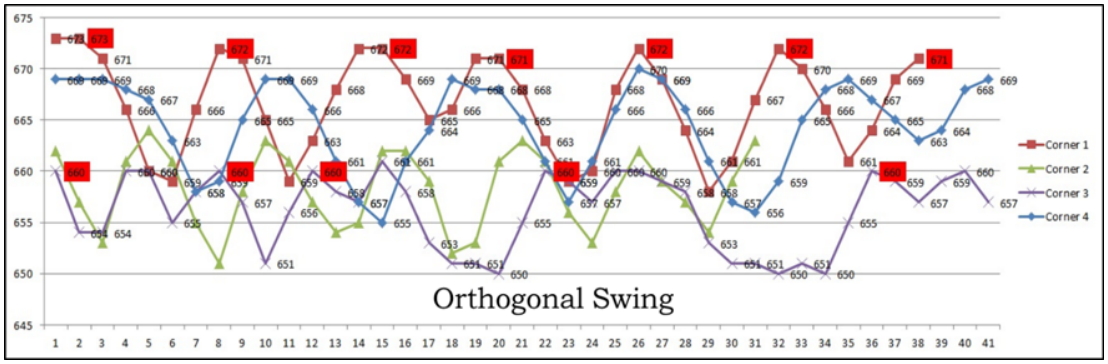
(a)

(b)

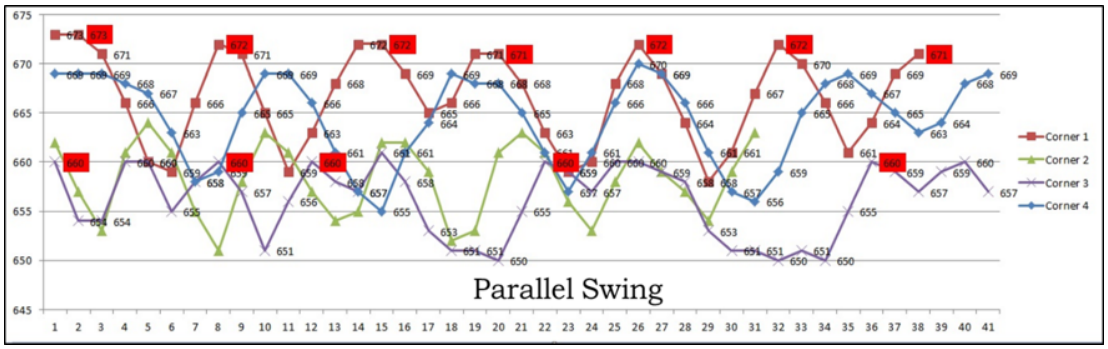
Fig. 54 (a) Topography of the roadway is measured using a total station to determine its highest and lowest points; (b) Position of the demonstrator at the 4 boundary locations in the road scene.

Method	Height of camera	Difference
Manually estimated	2.40 m	
Total station	2.33 m	
Pole	2.32 m	0.0081 m

Table 19 Results of the pole versus total station methods for height accuracy.



(a)



(b)

Fig. 55 Demonstrates the pole swung (a) orthogonal and; (b) parallel to the camera, with a difference of ±12 pixels on the image when manual levelling is performed.

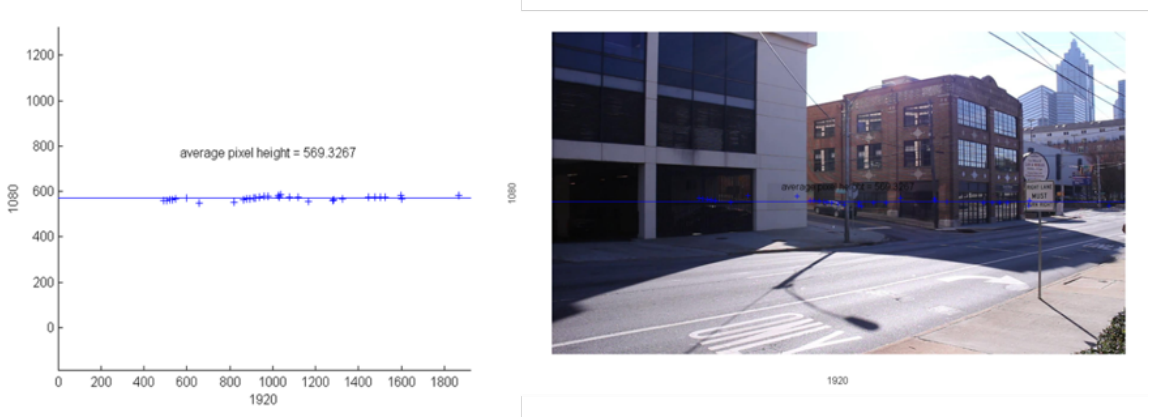


Fig. 56 Pole points, as projected onto image plane.

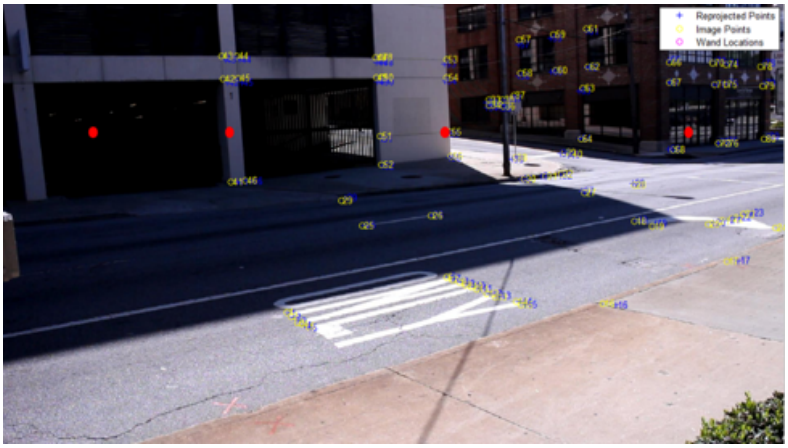


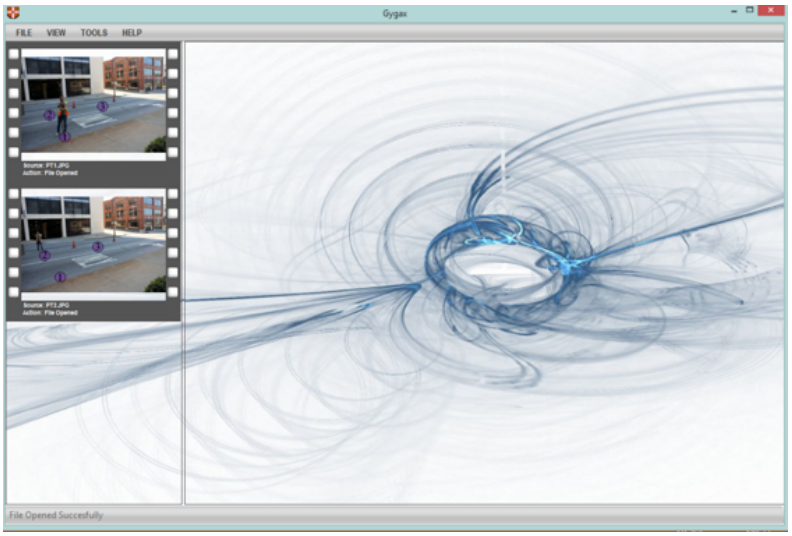
Fig. 57 2D points of the pole locations were back-projected, using total station 3D calibration points as a comparison, to find the height of the camera. Red dots represent the top locations of the poles. The height of the camera resulted in a T_z (translation along the x-axis) of 2.32 m.

at the respective marker locations (1) and (2) as shown in Figure 58. The user selects the pole tip marker in the image. The prototype records the pixel locations of the points and compares their y-pixel values. If the y-pixel values do not match, the prototype instructs the user to adjust the camera by a specified amount. The same procedure is carried out for θ_{roll} of the camera at marker locations (2) and (3). This process may require a series of iterations; this process may require a series of iterations, and takes 15 to 60 minutes. The process is designed to allow those with no prior experience to perform the calibration process. The process can be performed by a single person; however, two are recommended. In this scenario, one person would handle the software, while the other would continually reposition the pole in its respective location, allowing for maximal efficiency in this crucial component of the setup process.

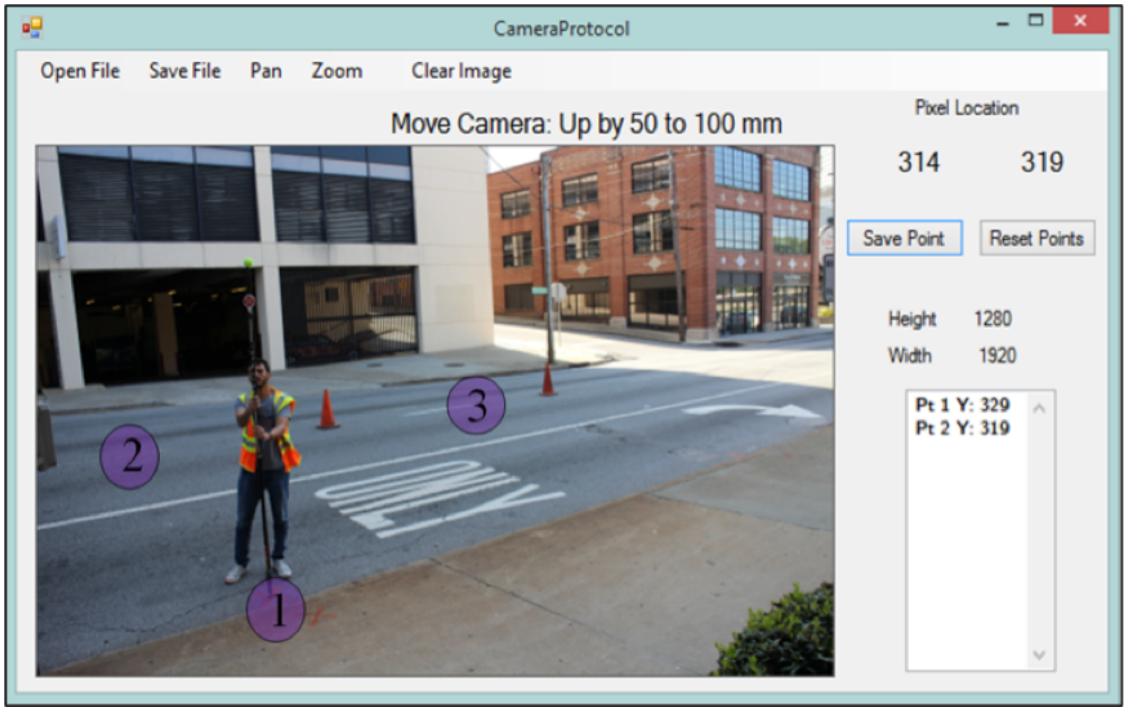
In Figure 58b is shown a screenshot of the prototype at Points 1 and 2 (i.e. marker locations (1) and (2)), together with their respective y-values. The prototype compares the differences in y-pixel values: 329 and 319, and instructed the user to move the camera vertically upwards by 50–100 mm. When the two points arrive at the same y-pixel value, the camera is assumed to be at the correct height for OH detection.

5.4.4 Height and detection accuracy (Over-height plane)

The first component of the experiment is performed 16 times to validate the installation procedure; a total station is used as ground truth data. A 'sanity check' is performed after each experiment to ensure the installation procedure's accuracy. This check consisted of capturing three undistorted photos at marker locations (1), (2) and (3). If the three world points projected onto the image view with the same corresponding y-pixel values, this would confirm that the camera is set at the correct height representing the OH plane.



(a)



(b)

Fig. 58 (a) Prototype camera installation tool to help users locate the over-height plane; (b) Shows the required camera adjustments to align the OH plane based on the two points extracted.

Height attempts by number	Height of OH plane (Pole Method)	Height of OH plane (Total Station)	Diff (mm)	Error (mm)
1	1784	1782	2	2
2	1803	1807	-4	4
3	1810	1810	0	0
4	1756	1755	1	1
5	1880	1884	-4	4
6	1768	1769	-1	1
7	1813	1819	-6	6
8	1800	1797	3	3
9	1756	1760	-4	4
10	1791	1794	-3	3
11	1821	1823	-2	2
12	1981	1983	-2	2
13	1795	1791	4	4
14	1897	1896	1	1
15	1765	1766	-1	1
16	2319	2327	-8	8
			Average Error:	**2.88**

Table 20 Comparing the height accuracies of the pole versus total station methods in millimeters.

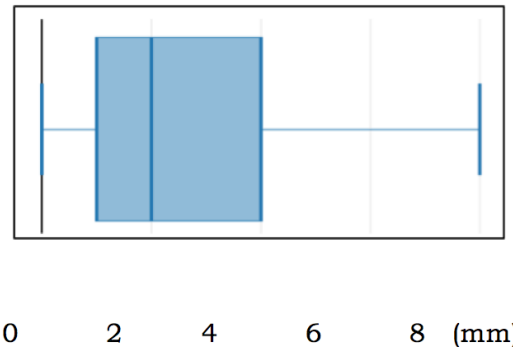


Fig. 59 Boxplot data comparing the error of the pole method against the total station method (validation data). The height error median is 2mm with an upper limit of 8 mm.

Two evaluations measured the height accuracy of the OH plane; (1) via pole method, and (2) via total station method. The ground truth data for method (1) is obtained by manual measurement, and for method (2) a total station is used to validate the height. Three points were measured for each of the experiments. The height accuracies are summarised in Table 20 and Figure 59, respectively. The two methods yielded an overall error of ±2.88mm.

5.5 Discussion and conclusions

This chapter has outlined a holistic solution to the overall problem of OHVS management, specifically contributing a novel approach to the prevention aspect. The method models an active laser sheet using passive vision methods, which constitutes a major improvement to the laser beam method in terms of practical viability. The chapter outlines both the installation and camera configuration procedure, underscoring how our new method is based on a simple geometric principle: the OH plane appears as a line on a 2D image given a camera mounted at the height of the bridge clearance and correctly calibrated. Any vehicle exceeding the OH line in the image view is deemed OH. Evaluation of the system resulted in a height estimation accuracy of ±2.88 mm, comfortably inside the target accuracy of ±5.00 cm.

The calibration process shows that the pole method can achieve near-millimetre accuracy when compared with ground truth data. We expect height accuracy error to be ±2.88 mm and the effect of the error to be ±0.04 mm per pixel in the real world (on the assumption that the calibration steps have been carried out as described above). On average, an OH vehicle was present in the scene 6.56% of the time during a period of 150 minutes of video data. The OH detection method was tested under ideal conditions: sunny, non-windy weather conditions resulting in an overall detection accuracy of 99.9% with a false positive rate of 0.1%. Two special cases were detected, in which a truck carrying a ladder and pole exceeded the OH plane in the image view, activating a warning. Although the consequences may have been less damaging than an entire truck striking the bridge, these instances meet the OH criteria of an OH vehicle and were therefore treated as true positives, with no negative consequences for accuracy evaluation.

As mentioned, the method has no need of estimating the exact height of OH vehicles. Rather, the method advocates a simpler approach returning one of two possible outcomes (OH / non-OH). The calibration process (setting the OH plane) is tuned to each specific low bridge plus roadway, but the performance of the algorithm is optimised for any site, given the same camera specifications, therefore the parameters for the detection algorithm are not scene-dependent. The calibration process takes less than 60 minutes to perform, and need not be repeated except in case of hardware damage.

The camera installation requires a bracket to be installed on an existing pole upstream of the low bridge as well as access to power and a processing unit, thus requiring a professional electrician. The calibration process can take between 15 and 60 minutes including

adjustment time. The camera setup is a permanent installation intended to remain in service for many years (5-15 years depending on camera life), hence the initial time investment is comparatively minor. Leading laser-based systems require permanent infrastructure installation requiring permit approvals, sub-contracting teams, engineers, planners, designers, road closures, road cuts, etc. The vision-based system does not require any of the above, which saves infrastructure owners significant upfront costs.

The method performed as expected according to the predictions of the camera modelling (i.e. camera height and orientation), with an overall height error of ±2.88 mm. The boxplot above shows the one-sided error, with a median height error of 2.00 mm and a maximal height error of 8 mm. All else being equal, the total station is to be preferred; however, total stations are expensive and require specialised training to operate. We thus compared the accuracy of the pole method to that of the total station method to determine whether the accuracy provided by the pole method is acceptable without having to purchase expensive equipment to determine the OH plane. The results demonstrated comparable accuracy between the pole installation and the total station method.

During the evaluation of the system, there were instances when the detection algorithm encountered operational issues such as swaying of the streetlight pole in the horizontal (x) and lateral (y) directions. The detection algorithm is unable to handle drastic pixel changes in windy conditions which contributed to the two false positive detections. Swaying in the horizontal axis have minimal effects on the OH line; however, if lateral displacements occur, it may offset the OH line in the image view and compromise the system accuracy. Further testing is required under more rigorous weather conditions to prove valuable in a real-world scenario. Chapter 6 thus extends the study by attempting to optimised the OH detection algorithm parameters under variable weather conditions.

Chapter 6

Optimising detection parameters

6.1 Introduction

In road traffic scenes, the background is generally hypothesised to be static for convenience. This, however, is not often the case in real-world traffic scenarios. Often, vision-based systems encounter obstacles such as moving trees, wind, rain and illumination changes caused by weather conditions. Outdoor cameras must also deal with gradual illumination changes caused by sunlight or clouds. Any vision-based OHVD system must be able to accommodate such changeability in order to minimise false positive detections. No vision-based OHVD systems are currently employed in practice. However, similar detection systems exist in the realm of traffic monitoring and enforcement. Motion detection systems fall under the domain of video surveillance, where security concerns are continually growing in importance (Lavanya, 2014). Motion detection plays a fundamental role in any object tracking or recognition algorithm.

Motion detection uses background subtraction, an image processing and computer vision technique which compares a given image to a reference image, specifically comparing white pixel intensity values with a threshold value to detect motion change. Motion detection focuses on an *area of motion* within a scene, also known as the ROI, which concept naturally translates well to the OHVD problem. If motion is detected within the ROI, a positive outcome is registered with respect to the binary classification task at hand. As mentioned, detection of motion within the ROI is not straightforward, due to changes in illumination and other confounding factors (Martinez-Martin & del Pobil, 2012). Such challenges apply equally to other applications of traffic monitoring systems, including flow monitoring, vehicle counting and speed cameras.

Cameras are used both in traffic flow surveillance systems, used in congestion management, and as enforcement devices, in systems commonly seen mounted beside roadways; these latter detect traffic violations including speeding vehicles, red light running, and unauthorised vehicles in bus lanes or congestion charge areas. These systems are typically linked with automated punitive ticketing systems. They are also housed in units designed

to protect against rain, hot and cold weather, dust, vibrations and vandalism. Infra-red cameras are used to detect violating vehicles in night-time conditions. Chapter 5 aims to answer:

> **Aim 2:** Determine the optimised image initialisation parameters to accurately detect OHV.

The chapter evaluates the system under adverse weather conditions by optimising the camera's initialisation parameters. The chapter is organised as follows: Section 6.2 describes the methods used to handle illumination and adverse weather changes. Section 6.3 introduces the optimisation technique used to find the ideal parameters for OHVD. Section 6.4 evaluates the parameters using a 1.9% generality dataset and Section 6.5 concludes with an overall evaluation of the performance of the algorithm.

6.2 Background

Stationary cameras are used to monitor activities at specific areas of interest for traffic video surveillance. In the case of outdoor OHVD, accuracy is largely contingent on successful handling of illumination changes and variable weather conditions. This can be effective using a background subtraction method to adapt to variable conditions.

Background subtraction is a motion detection algorithm that begins with the segmentation of the moving object (foreground image) from the background. Further processing is then required to better understand moving objects in scene events (Elgammal et al., 2002). The foreground image is formed of white pixels which reflect motion. Background subtraction uses a frame differencing approach that takes a certain background image (for OHVD, the road scene with no vehicles present), here denoted by $Ibackground$, and compares it with the frame obtained at time t denoted as $Icurrent(t)$. The frame differencing approach can then be expressed as:

$$P[Iforeground(t)] = P[Icurrent(t)] - P[Ibackground] \tag{10}$$

which segments moving vehicles into blobs at each pixel in image $Icurrent(t)$ with the corresponding pixels at the same position on the background image $Ibackground$. The foreground image $Iforeground(t)$ displays intensity changes at those pixel locations which have changed from the background image. To improve motion segmentation, a thresholding technique is commonly used to replace each pixel in an image with a black pixel if the white pixel intensity $Iforeground(t)$ is less than some fixed threshold value T (that is, $Iforeground(t) < T$). Therefore, in the difference image, the pixel intensities are compared against the threshold value expressed as:

$$|P[Iforeground(t)] - P[Icurrent(t)]| > Threshold\ value \tag{11}$$

One of the many challenges of the thresholding process is that we only consider the intensity of the pixel without considering relationships between pixels, e.g. identifying whether the pixels are contiguous or have features in common. Hence, we next briefly describe some improvements on the standard background subtraction process.

Three common approaches to background subtraction are the Gaussian Mixture Model (GMM) (Gorur & Amrutur, 2011; Pathan, Chauhan, & Kathiriya, 2016), kernel density estimation (Elgammal et al., 2002) and codebook (Kim et al., 2005) methods. The Gaussian Mixture Model (GMM) was introduced by Stauffer & Grimson (1999) for real-time outdoor tracking of objects. The algorithm assumes that the background is visible more frequently than the foreground, and that the expected value of the Gaussian distribution corresponds to the larger clusters. GMMs were improved upon by Zivkovic & Van Der Heijden (2006), who used a kernel estimation approach to automatically adapt to the scenes using a non-parametric approach suitable for static cameras. However, in dynamic and fast variating scenes, the background cannot be accurately modelled with a set of Gaussians (Vargas et al., 2010). An alternative method is the codebook background subtraction method.

The codebook models illumination changes at each pixel, and calculates the variation that occurs between the foreground and background (Kim et al., 2005). The method adapts well to novel scene conditions such as illumination, weather conditions and shadows, and hence appears ideal for OHVD. However, the videos relies on choosing sufficient threshold parameters to minimise the amount of noise. An optimal threshold value is thus required.

The threshold value can crucially affect the performance of later processing steps (Martinez-Martin & del Pobil, 2012). For instance, with a more relaxed threshold value, any slight illumination change can be detected by the algorithm therefore increasing the noise within the ROI; increased noise leads to false positive detections. With a more stringent threshold, conversely, risks the potential of losing value information of objects (especially near the boundaries of the region), leading to false negatives in an OHVD situation. To experiment with the threshold value, Unzueta et al. (2012) used a trial and error approach by lowering the threshold until vehicle shapes have sufficiently limited surrounding noise, a strategy that can also be applied to OHVD. Because trial-and-error threshold estimation, whereby filter pixel response values are examined using visual analysis, is inherently subjective, an iterative optimisation process approach is necessary here to determine the optimal value for accurate OHVD.

6.3 Proposed approach

The revised OHVD process is schematised in Figure 60. The camera initialisation was carried over from the experiments described in Chapter 5 of this thesis. The extension of the framework, as mentioned, involves determining the optimal parameters by using an iterative optimisation process. The purpose of the optimisation process is to determine the optimal parameters for: 1) $cv(threshold)$, filter pixel response value, i.e. threshold value

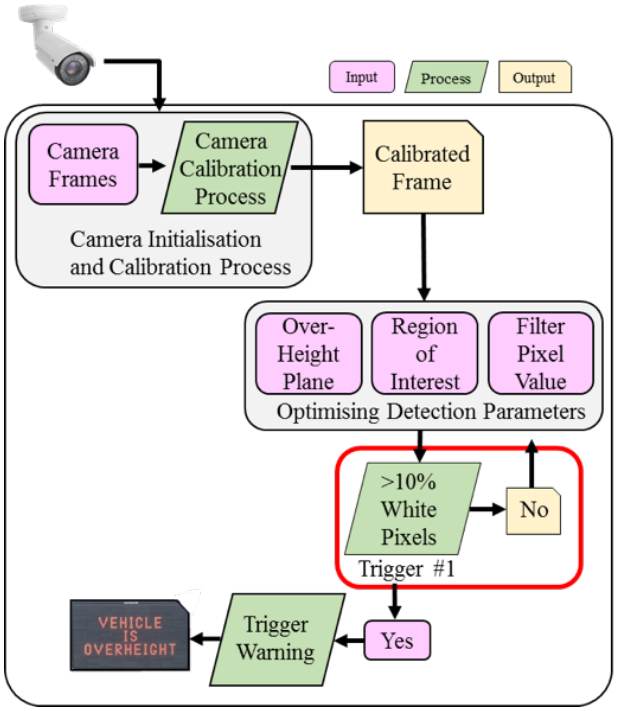


Fig. 60 Proposed approach for optimising sensitivity detection process for over-height vehicle detection using a trigger-based approach.

using a background differencing algorithm and 2) $cv(vert)$, vertical pixel height, i.e. the vertical region size above the OH plane to detect OH vehicles. These parameters will be used as the initialisation parameters for future OH vision-based systems.

The first step in the iterative optimisation process is initialising the OH plane in the image. The optimisation process uses three input parameters: $cv(threshold)$, $cv(vert)$ and $dv(horz)$. Next, a video containing all positively and negatively classified frames is passed to the algorithm. Each video frame is analysed using a binary classification (0 = non-OH and 1 = OH) and the entire dataset is scored using precision and recall metrics. The input parameters with the highest output values are then used for the next iteration. This process is performed until the algorithm converges on an optimal set of input parameters. Table 21 shows the initial settings and parameter limits relevant to optimisation: 1) the dependent variable $dv(horz)$ horizontal axis (x) at 1920 pixels (to maximise the camera's field of view), and 2) the two control variables $cv(threshold)$ and $cv(vert)$.

The purpose of the filter pixel response value is to optimise detection of moving objects within a specified ROI (area in which OH vehicles are present). The ideal filter pixel response value can be found by comparing the new image frames to a background model with various values, i.e. by trial and error. The background filter pixel response value use pixel values from 0–255; a grayscale image was used for classification as it contains sufficient pixel information for binary classification, given dataset (X_i, Y_i) for $i = 1,\ldots,N; t = 0,\ldots,255$ with $X_i \in \mathbb{R}d$ and $Y_i \in \{0,1\}$, classifier $f(x)$ such that

Control variables	Units	Lower limit	Upper limit
$cv(threshold)$	intensity value	0	255
$cv(vert)$	pixels	1	275
Dependent variables	Units	Lower limit	Upper limit
$dv(horz)$	pixels	1	1920

Table 21 Initial settings of control and dependent parameter limits for optimisation procedure.

$$f(x)\begin{cases} \geq t\ Y_i = 1 : OH\ vehicle \\ < t\ Y_i = 0 : not\ OH\ vehicle \end{cases} \tag{12}$$

The ROI, with window size $cv(vert)xdv(horz)$, is a box situated above the OH plane. The dimensional parameters are passed through the iterative optimisation process to find the ideal region size to accurately detect OH vehicles. This process is performed by inputting the $dv(horz)$ as a dependent variable at 1920 pixels wide to utilise the entire horizontal field of view while varying the vertical size of the window frame, $cv(vert)$. The vertical size of the window is varied along the y-axis of the image frame. If the window is too large, more noise will be captured however, if the window is undersized, important information may be lost. Therefore, to find the ideal window size for detection, the control variable $cv(vert)$ is iteratively optimised.

White pixel values with intensity close to 255 are used as trigger points to determine if there is motion within the ROI. If motion is detected, the algorithm calculates the number of white pixels present in each region of the image, and returns a percentage value. If the percentage value is above or equal to a trigger point value; this event will be flagged as a positive OH instance.

6.4 Experiment and results

This section details the experiments designed to test the various possible parameter settings elucidated above. The implementation took place in Atlanta, Georgia (USA) on a collector roadway with 4 lanes of traffic in sunny, cloudy and rainy weather conditions. The events are mutually exclusive as they were recorded on separate days ensuring that conditions were ideal for testing. A Canon EOS M camera was used to capture 150 minutes of video data (1920 x 1080 resolution) at 30 fps.

The dataset uses a generality of 1.9% positive retrieval rate. The generality represents the expected retrieval rate of randomly selecting a positive frame (i.e. one containing an OH vehicle). Table 22 shows the sample size calculation based on the generality calculation by Huijsmans & Sebe (2005). The generality of the dataset is calculated by taking the number

Sample size calculation	Number of frames
Negative frames	190303
Positive frames	3661
Total frames	193964
Generality frames	3661
	193964
Expected random retrieval rate	1.9%

Table 22 Sample size generality calculation.
Note: Negative frames = number of irrelevant items for a particular query = embedding size;
Positive frames = number of relevant items for a particular query = relevant class size;
Total frames = total number of items in the ranked database = database size.

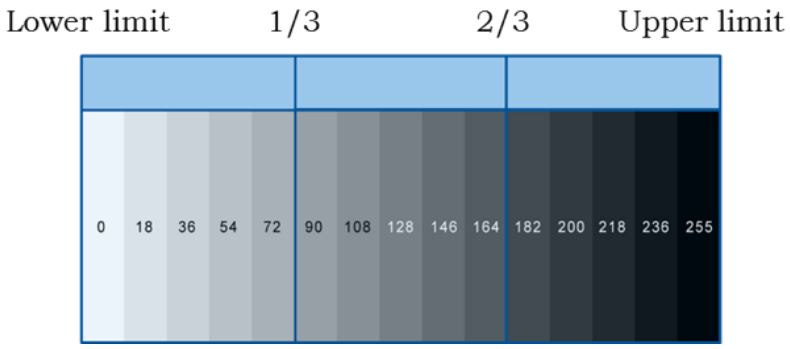


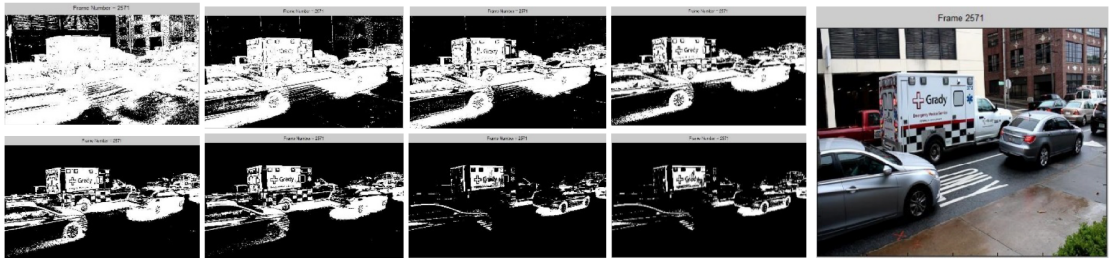
Fig. 61 Filter pixel response value: for the optimisation process, the $cv(vert)$ is divided into three quadrants ranging from a lower to upper limit.

of positive frames and dividing this by the total number of frames. In combination, the total frames are as in Table 22.

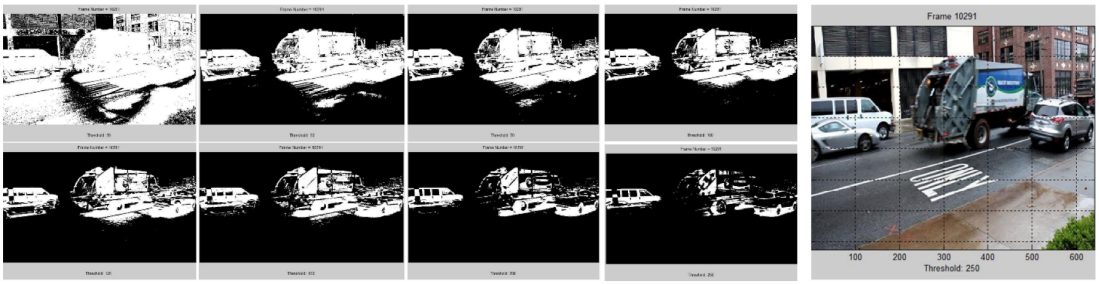
6.4.1 Filter pixel response value

Figure 61 illustrates the upper and lower limits of the control variables used as inputs to iterative optimisation process. For example, $cv(threshold)$ has a maximum range from 0–255. Therefore, for the optimisation process the range is divided into three equal quadrants used as input parameters. This would equate to four input values: 0, 85, 170 and 255. The optimisation process uses input values, 0 for example, and parameters: $cv(vert)$ and $dv(horz)$ as its starting iteration. The output is a score using precision and recall metrics to rate its performance. The set of input values with the highest output scores are then used for the subsequent iteration. The goal is for each of the control variables to converge to a value.

The input values ranged from 0 (relaxed threshold) to 255 (restricted threshold). The images were grayscale, with each pixel representing a single intensity value from 0 (white) to 255 (black). In Figure 62, scenarios 1 and 2 visualise differences at each of the filter pixel response settings.



(a) Scenario 1



(b) Scenario 2

Fig. 62 Two traffic scenarios depicting a series of filter pixel response values varying from 0 to 255 intensity values.

6.4.2 Region of interest

OH vehicles are detected within the ROI detailed earlier in camera calibration, but we want to know the vertical extent optimal for accurate OH detection. As the window size increases, the amount of noise captured can vary tremendously. Figure 63 shows an example of the ROI with a varying window size ranging from 1 to 275 x 128. The four options show the maximum height $cv(vert)$, 275 divided into three equal quadrants. This equates to input values 1, 63, 128, and 275. The input values, for example starting with value 1 was paired with a $cv(threshold)$ and $dv(horz)$ value to start the first iteration. The process is the same as in Section 6.4.1 where the input values was scored using precision and recall metrics and the highest scored set of inputs was used in the subsequent iteration. This process was iterated until the highest scored values was known in each of the three parameters.

For each combination of control variable settings, white pixel values are used to determine whether there is motion within the ROI. If motion was detected, the number of white pixels present within each region is calculated. It is correct that there would always be many more white pixels in the larger ROI than in the smaller ROI, however, the aim of this experiment determines the ideal ROI (window size) that would minimise the greatest amount of noise, at the same time as large enough to detect an OHV. Trigger point values were spaced at 10% intervals between 10 to 100%. If this value is above a specified threshold t, we classified this as a positive instance.

Fig. 63 The vertical height of the region of interest is 275 pixels tall, the optimisation process iteratively analyses the height to find the optimum vertical value.

6.4.3 Results

Two performance metrics were used to evaluate the detection algorithm: 1) precision and recall metrics to evaluate the performances of the control variables and 2) receiving operating characteristic curve (ROC) representing the performance of the algorithm. Table 23 shows the precision & recall metrics used after each of the optimisation iterations, where the positive class (Y) = 1 and negative class ($\hat{Y}$) = 0. $\hat{Y}$ is the estimate of the true class label Y.

'Recall' represents returned true positives as a proportion of real-world positives, equivalent to true positives as a proportion of true positives plus false negatives. 'Precision' represents the proportion of returned positives which were true positives, or in other words, the correctly classified proportion of returned positives. The acceptable recall rate for the system was set to 0.950 – 1.000, or < 5% missed OHV. The detection rate (DR) is the number of OH vehicles detected divided by the total number of OH vehicles. The false alarm rate (FAR) is the number of false alarms divided by all messages issued by the system, where

$$precision = tp/(tp + fp) = P(Y = 1|\hat{Y} = 1) \tag{13}$$

$$recall = sensitivity = tp/(tp + fn) = P(\hat{Y} = 1|Y = 1) \tag{14}$$

$$specificity = P(\hat{Y} = 0|Y = 0) \tag{15}$$

Figure 64 shows the results of the ROC curve of 83.3% (area under the curve) at the final optimisation iteration. Figure 65 shows the results of the iterative optimisation process.

Precision & Recall	Relevant	Nonrelevant
retrieved	true positives (*tp*)	false positives (*fp*)
not retrieved	false negatives (*fn*)	true negatives (*tn*)

Table 23 Precision and recall metrics.

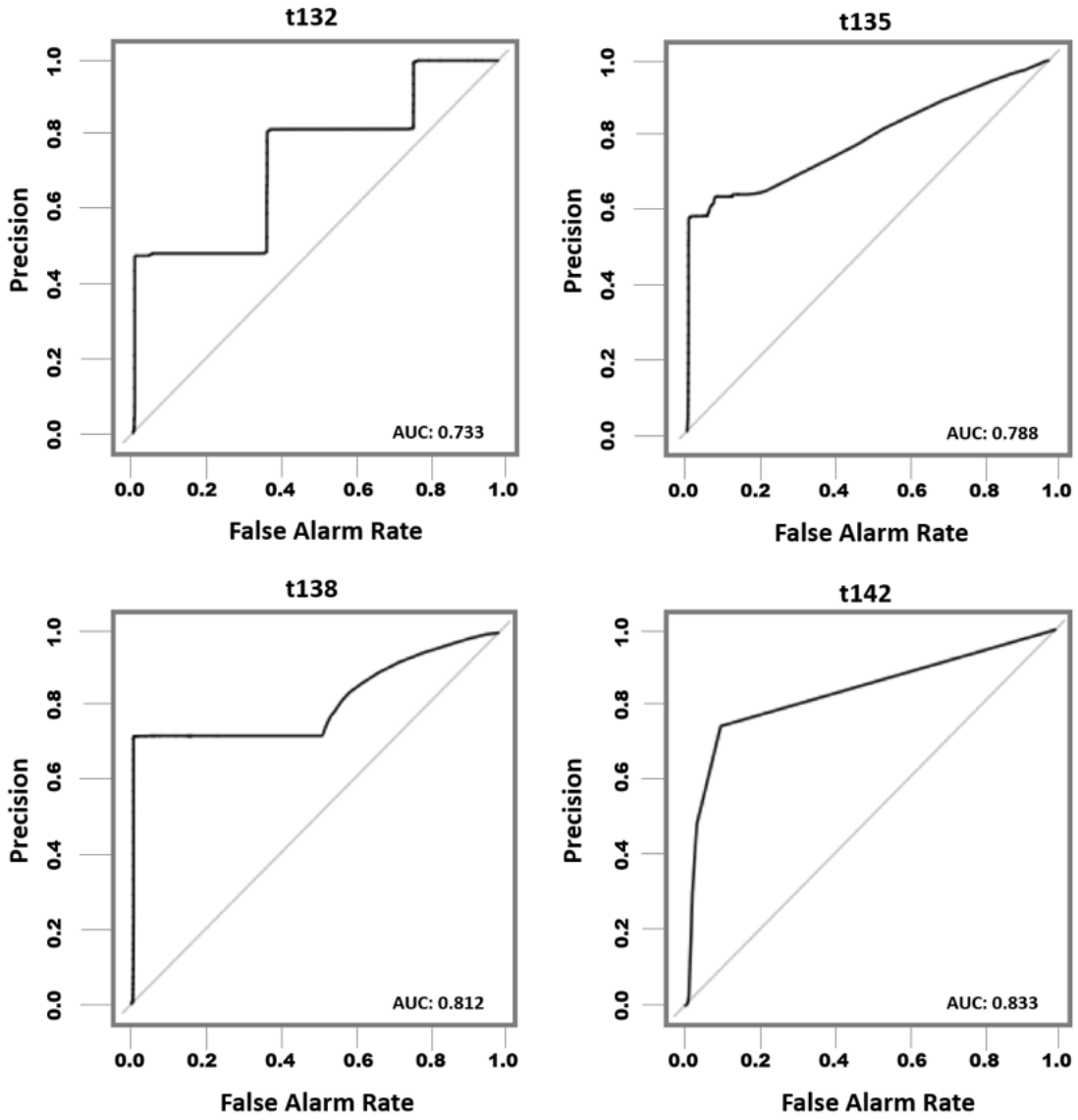


Fig. 64 Receiving operating characteristic (ROC) curve results, performance of classifier at final optimisation iteration: $cv(vert) = 142 * dv(horz) = 1920$.

cv(vert_ROI)	*cv*(threshold)		iteration_1 trigger points 10%	20%	30%	40%	50%	*cv*(vert_ROI)	*cv*(threshold)	iteration_2 trigger points 10%	20%	30%	*cv*(vert_ROI)	*cv*(threshold)	iteration_3 trigger points 10%	20%	30%	*cv*(vert_ROI)	*cv*(threshold)	iteration_4 trigger points 10%	20%
1x1920	0	P	0.223	0.227	0.227	0.227	0.228	63x1920	85	0.462	0.469	0.487	63x1920	113	0.482	0.545	0.566	63x1920	132	0.524	0.603
		R	1.000	1.000	1.000	0.456	0.426			1.000	1.000	0.596			1.000	1.000	0.710			1.000	0.951
	85	P	0.290	0.290	0.301	0.306	1.000		113	0.482	**0.545**	0.566		123	0.516	0.517	0.517		135	0.527	0.603
		R	1.000	1.000	0.665	0.654	0.456			1.000	**1.000**	0.710			1.000	1.000	0.694			1.000	1.000
	170	P	0.366	0.366	0.703	0.710	0.722		142	0.532	**0.628**	0.693		132	0.524	0.603	0.628		138	0.530	0.615
		R	1.000	1.000	0.752	0.751	0.564			1.000	**1.000**	0.858			1.000	0.951	0.861			1.000	0.951
	255	P	0.250	0.256	0.256	1.000	1.000		170	0.499	0.498	0.750		142	0.532	**0.628**	0.693		142	0.532	0.628
		R	1.000	0.740	0.732	0.634	0.503			1.000	1.000	0.817			1.000	**1.000**	0.858			1.000	1.000
63x1920	0	P	0.290	0.290	0.301	0.306	0.457	85x1920	85	0.328	0.329	0.330	70x1920	113	0.530	0.530	0.530	65x1920	132	0.564	0.566
		R	1.000	1.000	1.000	1.000	0.992			1.000	1.000	0.822			1.000	1.000	0.822			1.000	1.000
	85	P	0.462	0.469	0.487	0.515	0.549		113	0.426	0.429	0.509		123	0.584	0.584	0.585		135	0.597	0.597
		R	1.000	1.000	0.596	0.565	0.565			1.000	1.000	0.654			1.000	1.000	0.654			1.000	1.000
	170	P	0.499	**0.498**	0.750	0.749	1.000		142	0.522	**0.522**	0.618		132	0.644	**0.664**	0.689		138	0.612	0.617
		R	1.000	**1.000**	0.817	0.814	0.672			1.000	**1.000**	0.817			1.000	**1.000**	0.817			1.000	1.000
	255	P	0.285	0.290	0.290	1.000	1.000		170	0.465	0.466	0.532		142	**0.689**	0.691	0.690		142	0.619	0.619
		R	1.000	0.636	0.594	0.312	0.312			1.000	0.422	0.357			**1.000**	0.521	0.520			1.000	0.521
128x1920	0	P	0.223	0.223	0.223	0.224	0.290	107x1920	85	0.322	0.322	0.331	78x1920	113	0.530	0.531	0.515	67x1920	132	0.608	0.641
		R	1.000	1.000	1.000	1.000	0.462			1.000	1.000	0.517			1.000	1.000	0.517			1.000	1.000
	85	P	0.492	0.492	0.521	0.531	0.546		113	0.448	0.449	0.535		123	0.540	0.540	0.561		135	0.628	0.643
		R	1.000	1.000	0.762	0.765	0.550			1.000	1.000	0.789			1.000	1.000	0.789			1.000	1.000
	170	P	0.501	**0.501**	0.764	0.769	1.000		142	0.447	0.447	0.477		132	0.632	0.631	0.514		138	0.650	0.655
		R	1.000	**1.000**	0.765	0.761	0.754			1.000	1.000	0.617			1.000	1.000	0.602			1.000	1.000
	255	P	0.241	0.241	0.284	0.324	1.000		170	0.380	0.382	0.458		142	0.621	0.540	0.521		142	0.682	0.686
		R	1.000	0.732	0.616	0.524	0.457			1.000	0.716	0.595			1.000	0.721	0.605			1.000	0.721
275x1920	0	P	0.196	0.195	0.201	0.201	0.237	128x1920	85	0.492	0.492	0.521	85x1920	113	0.426	0.429	0.509	70x1920	132	0.644	0.664
		R	1.000	1.000	1.000	1.000	0.457			1.000	1.000	0.762			1.000	1.000	0.654			1.000	1.000
	85	P	0.436	0.436	0.446	0.502	0.502		113	0.451	0.451	0.452		123	0.549	0.549	0.550		135	0.681	0.681
		R	1.000	1.000	0.613	0.613	0.613			1.000	1.000	0.722			1.000	1.000	0.695			1.000	1.000
	170	P	0.456	0.457	0.616	0.628	0.628		142	0.451	0.451	0.452		132	0.530	0.530	0.505		138	0.682	0.682
		R	1.000	1.000	0.537	0.537	0.537			1.000	1.000	0.537			1.000	1.000	0.525			1.000	1.000
	255	P	0.216	0.232	0.246	0.247	1.000		170	0.501	0.501	0.764		142	0.522	0.522	0.618		142	**0.689**	0.691
		R	1.000	0.316	0.316	0.316	0.307			1.000	1.000	0.765			1.000	1.000	0.817			**1.000**	0.521

Fig. 65 Iterations of optimisation process; final iteration resulting in recall 1.0000 and precision 0.689 highlighted in green.

Optimisation converged at a window size of 70 (±3) x 1920 pixels and a filter pixel response value of 142 (±5). The average precision value is 0.689 (sunny: 0.751; cloudy: 0.631; rainy: 0.685), with recall of 1.000 as shown in Fig. 64. The results showed that the minimum proportion of white pixel values ('trigger points') required to detect an OH vehicle is 10%.

6.5 Discussion and conclusions

In this section, an extended version of Chapter 5 is presented, which develops an optimised computer vision approach to OHVD under variable weather conditions. The system achieved a recall of 100% of vehicles and precision of 68.9%. The ROC curves resulted in a classification performance of 83.3%. Optimisation converged at a window size of 70 (±3) x 1920 pixels and filter pixel response value of 142 (±5). At the threshold value of 142, the ROC results showed a step variation, (as shown in Figure 63 under t132), arising from clustered observations of OHV with similar features and similarities in the background scenes. As we increased the filter threshold response value, values marginally improved, returning a classification performance of 83.3% at the final optimisation iteration.

In the experiment, only recall values between 0.95 and 1.0 were considered. Any values with more than 5% of missed OHV were discarded in subsequent iterations of the optimisation process, and so less than 5% of OHV were missed. Discarding input parameters where more than 5% of vehicles were missed allows the system to maintain high detection accuracy. Although the algorithm correctly classified 100% of OHV, the precision varied significantly according to experimental condition resulting in 31.1% of reported positives to be false alarms. The average precision value is 0.689 (sun: 0.751; cloudy: 0.631; rain: 0.685), with recall of 1.0. The average is calculated at only one pair of parameter settings (the optimal pair) not over the different sets of input values. In future work, the weather conditions can be added as a control variable and tested individually for its performance. One of the factors for false alarms is the result due to wind.

In the latter, the pole on which the camera is mounted is susceptible to lateral and horizontal sway. As the window size increased, more background noise due to this sway is encountered, leading to a sharp increase in false positive detections. Horizontal sway demonstrated only minimal effects on OH line projection to the 2D image; however, during windy gusts, more white pixels were captured within the ROI leading to triggering the algorithm. In severe windy conditions, this can be considerably a risk and therefore compromise the system accuracy.

Future development of this work must include systematic assessment of camera motion and stabilisation under variable weather conditions, in order to further minimise false positive detection instances. A further desirable extension, to the same end of improved precision, would include analysis of motion vectors of vehicle trajectories. The motion vectors of vehicle trajectories can be tracked to differentiate whether an OHV is present in the scene. Vehicle trajectories can be tracked over a series of video frames therefore

providing a more robust measure of classification. If the displacements do not behave (or move) as an expected vehicle should, the motion is regarded as noise. Chapter 7 explores further methods of vehicle displacements and video stabilisation to minimise the number of misclassifications with an overall system validation to conclude in Chapter 8.

Chapter 7

Minimising False Positive Detections

7.1 Introduction

OHVD systems using computer vision classification algorithms must avoid the risk of returning false positives due to windy conditions, in order to ensure the accuracy of OH warnings promulgated to drivers in real-time.

Chapter 6 of this thesis reveals that under ideal weather conditions (low wind, concomitantly, lack of precipitation / other occlusion), the distinction between OHV, wind and other noise is distinguishable by the number of feature points detected. Extrapolating from ideal conditions to *all* weather conditions, the hypothesis is that whenever the number of detected feature points exceeded a threshold value, for example, 250 features points detected, OHV would indeed have passed through the frame (i.e. the case would be a true positive); however, this proved incorrect. In windy weather conditions, the results proved to be inconclusive resulting in a decrease of performance by nearly 31.0%, all due to false alarms. Further refinement to and testing of the system is thus required. Therefore, the aim of this chapter is to:

> **Aim 3:** Minimise the number of false positive detections.

The chapter is divided into two parts:

> **Part 1** (Wind Analysis Check): What is the sampling and frame rate needed to accurately warn drivers in windy weather conditions?
>
> **Part 2** (Vehicle Displacement Check): What feature detector performs best to track vehicle displacement given the frame and sampling rates?

Part 1 performs a wind analysis check to determine whether the object detected has a constant and positive flow movement based on OHV characteristics and behaviours. The wind analysis check requires the ideal sampling and frame rate necessary in order to warn drivers adequately. Part 2 performs a final vehicle displacement check to determine whether

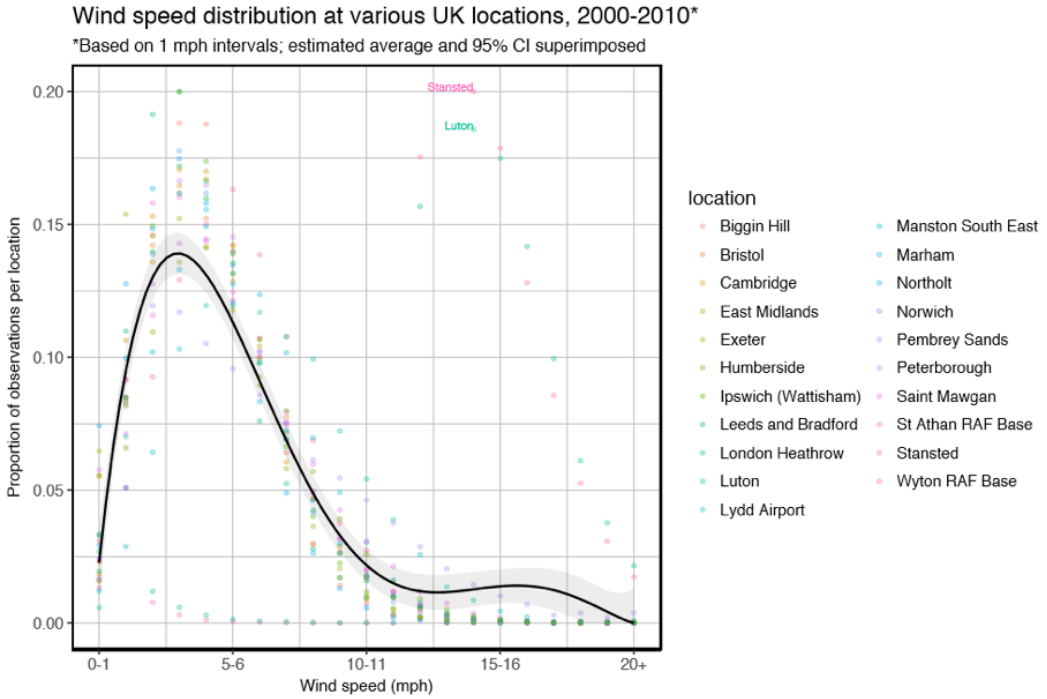


Fig. 66 UK wind speeds (2000 - 2010) in 1 mph intervals; estimated mean proportion per speed of all airport locations and 95% confidence interval superimposed.

the object is continuous over x number of frames to minimise any potential objects that may not be OHV. The vehicle displacement check compares the precision, recall and false positive rate using state-of-art feature detectors to determine the ideal method.

This chapter is organised as follows: Section 7.2 explores negative effects of wind in outdoor computer vision applications, and feature detection methods for vehicle detection. Section 7.3 steps us through the proposed approach to tackle windy camera motion and vehicle displacement checks. Section 7.4 reveals the results of the experiment and Section 7.5 concludes with a discussion and conclusions.

7.2 Background

Computer vision-based OHVD is a relatively new area of research, which has grown in popularity as infrastructure owners have sought more affordable methods of strike prevention. However, a reliable camera-based setup may prove challenging in windy conditions, hence understanding the characteristics of wind motion is crucial in designing an intelligent OH detection system. The National Weather Service UK defines 'windy' as speeds of 15 to 25 mph. According to Figure 66, the typical UK wind speeds averages ±3 mph. Camera installation locations will need to consider the wind levels in specific regions to maintain robustness during wind activity. The following sections describe our attempted employment video stabilisation to this end.

7.2.1 Video stabilisation methods

Camera stabilisation is a device that is used to securely fasten a camera in a manner that prevents or compensates for unwanted camera motion. In an OH context, although the camera is intended to be static and fastened to a secure bracket, severe wind can still cause motion in the system, due to the slenderness of the mounting pole and its height relative to the road plane. This problem can be addressed using feature- or intensity-based methods. Related methods to tackle this problem of windy camera frames can be explored using video stabilisation methods.

Computer vision image enhancement techniques for video stabilisation is the process of identifying and removing undesired image motion from video data. Shaky and blurry video data due to wind movements suffer from significant amounts of unexpected image motion caused by external weather conditions. The initial step in video stabilisation is global motion estimation followed by feature and intensity-based methods.

Global motion estimation is a vital step in the process of video stabilisation for OHVD. Huang et al. (2010) describe the process as 'what is happening in the frames and what motions are evident?' By understanding what is happening, the motions can be separated into two categories: 1) intentional motion (what we are trying to analyse) and 2) unwanted motion (camera jitteriness, jerkiness, background noise and wind motions). By removing the unwanted motion, we are then left with a stabilised frame. The use of video stabilisation techniques as part of a pre-processing stage can be used to bridge this concept to minimise the number of false positive detections invoked by the system. Methods for video stabilisation fall under two categories: 2D and 3D methods. Although 3D methods are more accurate, the method is more computationally complex and 2D methods are sufficiently robust to solve the problem of windy camera frames. A general survey of approaches to address this challenge of feature and intensity-based methods are presented.

7.2.2 Feature-based methods

This section covers the current feature-based extraction methods. Feature based methods has become an increasingly used method for detecting objects in a scene, ideal for OHVD which includes distinctive attributes such as edges and corners. The edge detection is one of the most practical and commonly used algorithms which treats edge detection as a signal processing problem that maximises the signal to noise ratio to provide good detection (Hocenski & Vasili, 2006). However, corners are mathematically the best features to track due to its difference in intensity values (Shi & Tomasi, 1994).

Corners are common features to track, due to their distinctive edges (as a sub-attributes of a corner) and distinguishable changes in intensity values at all vertexes. Corners are important features in the video as it contains high contract and high curvature points. During day time conditions, the corner information preserves a stable and coherent motion region (Jazayeri, Cai, Zheng, & Tuceryan, 2011). In Figure 67, each pixel p in the neighbourhood

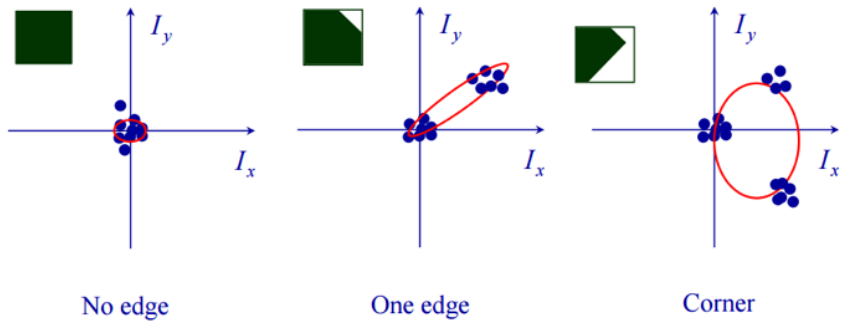


Fig. 67 Feature detection representing the major and minor axis of the elliptical approximation of the gradient vector distribution for no edge, one edge and corner (University of California at Santa Cruz, 2017).

of q, the gradient is $[I_x I_y]^2$ and C is the covariance matrix of all gradient vectors in q. In the same figure, the eigenvalues represent the major and minor axis of the elliptical approximation of the gradient vector distribution.

Corners are rotation-invariant, which means that if an image is rotated, the same corners can be found despite its orientation. The eigenvalue determines whether a region is a corner, edge or flat. When λ_1 and λ_2 are small, $|R|$ is small, therefore the region is flat. When $\lambda_1 > \lambda_2$ or vice versa, $R < 0$, therefore the region is an edge. When λ_1 and λ_2 are large, R is large and therefore, the region is a corner. Detected corners are often referred to as keypoints, otherwise as 'points of interest'.

In computer vision, the concept of keypoints have been used to address many problems in visual tracking, object recognition and 3D reconstruction. The method relies on the idea of focusing on a region of interest in the image to select interesting points known as feature points and to performs a local analysis around the points. An advantage of using the keypoints is that they permit matching even in the presence of clutter and occlusions and large scale and orientation changes (Sadeghi-Tehran, Clarke, & Angelov, 2014).

The common leaders in keypoints are known as *Scale Invariant Feature Transform* (SIFT) and *Speeded up Robust Feature* (SURF) (Lowe, 2004; Bay, Ess, Tuytelaars, & Van Gool, 2008); however, SURF outperforms the SIFT feature detector by computational speed (Dawood, Cappelle, El Najjar et al., 2012; Valgren & Lilienthal, 2010). These two methods are patented in the US and not free for commercial use. SURF is similar to SIFT, however the algorithm SURF pushes the SIFT algorithm further with box filters which approximates second order Gaussian derivatives using integral images. The advantage of box filters can be easily calculated using integral images and in parallel for various scales. Scale and location is determined using the determinant of Hessian Matrix. SURF has three main parts: 1) interest point detection, local neighbourhood description and matching. The detector is based on the Hessian matrix due to its good performance and computation time and accuracy. An interst point is similar to a feature point that refers to a point in the image which in general can be characterised as a clear, mathematically well-founded, position in image space. To detect

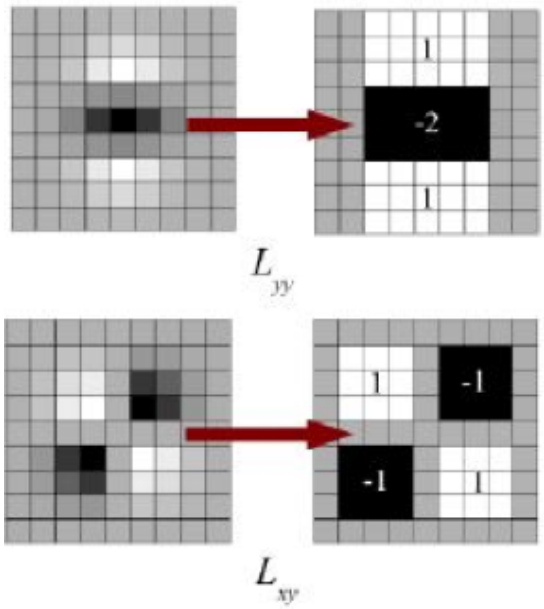


Fig. 68 Gaussian second order partial derivatives in y-direction and xy-directions using box filter approximations (Bay, Tuytelaars, & Gool, 2006).

interest points, SURF uses an integer approximation using the determinant of Hessian blob detector. Given a point $x = (x, y)$ in an image, I the Hessian matrix, $H(x, \sigma)$ in x at scale σ can be defined as

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \tag{16}$$

where $L_{xx}(x, \sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point x and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$, as shown in Figure 68.

Filtering the image with a box using integral images is much faster when the sum of the original image within the box is evaluated. The images are repeatedly smoothed with a Gaussian and sub-sampled to get the next higher level of the pyramid. In a Gaussian pyramid, images are weighted down using a Gaussian average (blur) and scaled down throughout the pyramid (multi-scale representation). Each pixel containing a local average that corresponds to a pixel neighbourhood on a lower level of the pyramid (Figure 69).

The scale space σ is divided into many octaves (series of response maps covering a doubling of scale). The lowest levels of the scale space are obtained from the output filters. The scale space is analysed by upscaling the filter size and the following layers are obtained by filtering the image with gradually larger masks. Non-maximum suppression (edge thinning technique) is applied to localise interest points in the image and over scales. The maxima of the determinant of the Hessian matrix are interpolated in the scale and image space.

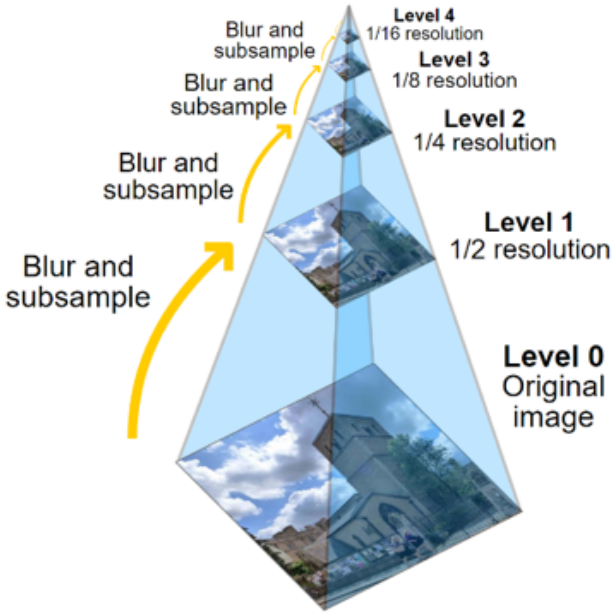


Fig. 69 A multiscale pyramid representation with 5 levels (Wikipedia, 2017).

The local neighbourhood feature descriptor is based on the sum of the Haar wavelet response around the points of interest. The Harr wavelet can be described as the region around the point, a box region is extracted (square), and centered on the interest point. The Haar wavelet is a sequence of rescaled squared-shaped functions which form a wavelet. The interest region is divided into smaller regions, 4x4 square sub-regions and the Haar wavelet responses are extracted at regularly spaced sample points. The responses are weighted with a Gaussian to ensure robustness for deformations, noise and translation. The matching is performed by comparing the descriptor obtained from the different images.

Interest points can be found at different scales with the various measures of masks calculated by:

$$\sigma approximate = Current filter size * \frac{Base filter scale}{Base filter size} \tag{17}$$

The results of the SURF detector and descriptor outperforms other feature detectors (SIFT, Harris) significantly and most importantly, the use of integral images improving the speed of detection. Other techniques include *Accelerated Segment Test* (FAST), *Binary Robust Invariant Scalable Keypoints* (BRISK) and Harris corner detector ((Harris & Stephens, 1988).

BRISK performs similarly to the state-of-the-art methods while dramatically more computationally efficient. A comprehensive evaluation comparing SIFT, SURF and BRISK reveals that BRISK's performance overall trumps the other feature detectors with a lower computational cost and at an order of magnitude faster than SURF in cases (Leutenegger, Chli, & Siegwart, 2011). The method generates keypoints from an image and keypoints are detected in octave layers of the image pyramid as shown in Figure 70.

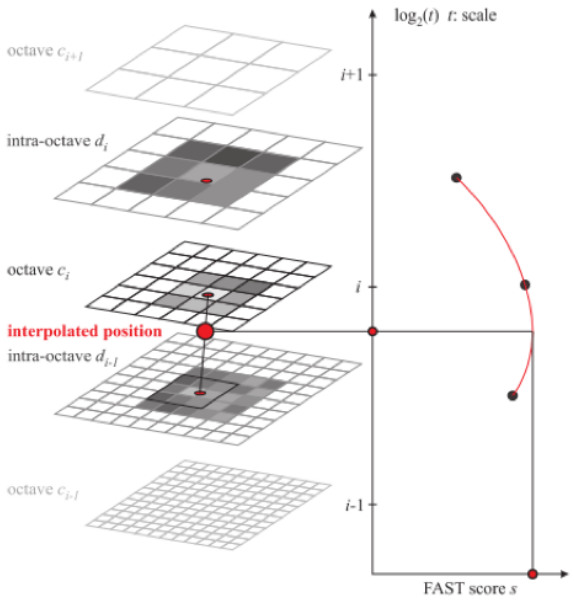


Fig. 70 Leutenegger et al. (2011) method of BRISK scale-space interest point detection using a keypoint, saliency maximum identified at octave c_i by analysing the 8 neighbouring saliency scores in c_i and corresponding scores-patches in the immediate neighbouring layers above and below. In the three layers of interest, the local saliency maximum is sub-pixel refined. The keypoint location is re-interpolated between the patch maxima closest to the determined scale.

The location and the scale of each of the keypoints are obtained using a quadratic function fitting method. FAST scores s are computed at each octave and intra-octave separately. The s scores are defined for each pixel as the maximum threshold for FAST detection (Features from Accelerated Segment Test) which still considers an image point a corner (Figure 71). Non-maximal suppression is performed on each octave and intra-octave such that s scores is maximal within a 3x3 neighbourhood.

A 2D quadratic function is fitted to the 3x3 patch surrounding the pixel and sub-pixel maximum is determined. The maxima's are then interpolated using a 1D quadratic function across scale space and the local maximum is chosen as the scale for the feature found.

The matching descriptor uses the hamming distance to compute all pairs between images. A threshold for matching relies on the number of bits able to be matched. The timing analysis compares the three state-of-art feature descriptors SIFT, SURF and BRISK (Figure 72) showing a faster BRISK computation time per point (ms) by 99.4% from SIFT and 91.3% from SURF. In Figure 73b, the time comparison in nanoseconds from the first and second image compares BRISK to SIFT by 69.4% faster and BRISK to SURF by 69.0% faster. The key to the speed lies in the application of FAST-based detector in combination with the assembly of a bit-string descriptor of each keypoint neighbourhood.

FAST feature detector is proposed by Edward Rosten and Tom Drummond in 2006 for identifying interest points in an image (i.e. corners). The algorithm uses a machine

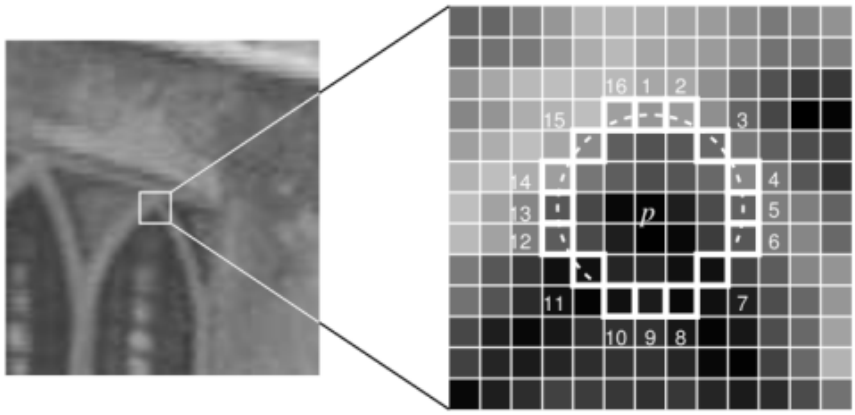


Fig. 71 The 16-point FAST detector is used, which requires 9 consecutive pixels which are sufficiently brighter or darker than the central pixel (Rosten & Drummond, 2005).

	SIFT	SURF	BRISK
Detection threshold	4.4	45700	67
Number of points	1851	1557	1051
Detection time [ms]	1611	107.9	17.20
Description time [ms]	9784	559.1	22.08
Total time [ms]	11395	667.0	39.28
Time per point (ms)	**6.156**	**0.4284**	**0.03737**

(a)

	SIFT	SURF	BRISK
Points in first image	1851	1557	1051
Points in second image	2347	1888	1385
Total time [ms]	291.6	194.6	29.92
Time per comparison [ns]	**67.12**	**66.20**	**20.55**

(b)

Fig. 72 Leutenegger et al. (2011) results for computation timing analysis (a) displaying the detection and extraction timings for the an image sequence (size 800 x 640 pixels) and (b) showing the matching timings for points in first and second image.

learning approach for use in real-time frame rate applications such as SLAM (simultaneous localisation and mapping).

The FAST algorithm works such that it selects a pixel p in the image and identifies it as 1) interest point or 2) non-interest point. The intensity of the pixel is represented by Ip and the appropriate threshold value is t. Consider Figure 71, a circle of 16 pixels around the pixel of interest. The pixel p is considered a corner if there exists a set of n contiguous pixels in the circle (of 16 pixels) with a brightness larger than $Ip + t$ or darkness than $Ip - t$. The white dash lines are represented as $n = 12$. The corner is validated by a high-speed test.

The test examines only four pixels at 1, 9, 5 and 13. The first test examines 1 and 9 to determine if the brightness or darkness thresholds apply. If true, then checks 5 and 13 are performed. If p is a corner, then three of the four pixels must be all brighter or darker. If neither of these are true, then p cannot be a corner. Paired with a machine learning approach, a set of images is selected for training. The FAST algorithm is performed to every image to find feature points (a vector representing the 16 pixels around each feature point). This process is performed for all images to get a feature vector p. Each pixel is analysed and classified as three types:

$$S_{p\to x}\begin{cases} d, & I_{p\to x} \le I_p - t & \text{(darker)} \\ s, & I_p - t < I_{p\to x} < I_p + t & \text{(similar)} \\ b, & I_p + t \le I_{p\to x} & \text{(brighter)} \end{cases} \tag{18}$$

Depending on each type, the feature vector p is subcategorised into P_d, P_s, P_b. A new boolean variable is created and each subset is queried to determine whether a corner exist and false if otherwise. This is a recursively process applied to all subsets until its entropy is zero. Non-maximal suppression is used to detect multiple interest points in adjacent locations. The edge thinning process computes a score function v for all detected features. v is the sum of absolute difference between p and the 16 pixel values. Two adjacent keypoint values are computed and the lower v value is discarded. The feature detector performs several times faster than other existing corner detectors at a lower cost on computation (Jiang, Xu, & Liu, 2013; Rosten et al., 2010). In the next section, optical flow is discussed as it uses the feature detectors to track useful points for OHVD.

7.2.3 Intensity-based methods

Optical flow is a widely used intensity-based method, which estimates the motion of image objects or discrete image displacements between two image frames at frames, I and I_{i+1} using the following assumptions:

1. The pixel intensities of an object do not change between consecutive frames.

2. Neighbouring pixels have similar motion.

Optical flow estimation is performed for 2D cases as the intensity $I(x,y,t)$ is shifted by d_x and d_y between the two image frames given as

$$I(x,y,t) = I(x+d_x, y+d_y, t+d_t) \tag{19}$$

where the x and y components represent the velocities of $I(x,y,t)$. Then, take the Taylor Series approximation to get the equation

$$f_x u + f_y v + f_t f_x u + f_y v + f_t = 0 \tag{20}$$

when simplified gives the optical flow equation

$$f_x = \frac{\partial f}{\partial x}; f_y = \frac{\partial f}{\partial y} \text{ (image gradients)} \tag{21}$$

$$u = \frac{d_x}{d_t}; v = \frac{d_y}{d_t} \tag{22}$$

The Lucas-Kanade method is another widely used differential method for optical flow estimation. The method assumes that the flow is essentially constant in a local neighbourhood of the pixel examined and using the least squares criterion, solves the basic optical flow equations for all pixels in that neighbourhood. Secondly, the method assumes that the displacement of the image contents between two nearby frames are small and constant within a neighbourhood of points under consideration using gradients weighted by an approximation to the second derivative of the image.

KLT feature tracker is one of the most widely used feature detectors today, noting that corners are good features to track (Jianbo Shi & Tomasi, 1994). KLT Feature Tracker is based on the two papers. The first paper describes the one-dimensional case. If d is the displacement between two images $I(x)$ and $G(x) = I(x+d)$ then the approximation is made that

$$I'(x) \approx \frac{I(x+d) - I(x)}{d} = \frac{G(x) - I(x)}{d} \tag{23}$$

so that

$$d \approx \frac{G(x) - I(x)}{I'(x)} \tag{24}$$

This approximation to the gradient of the image is only accurate if the displacement of the two images is small. The approximation d depends on x. The next step averages the estimates of d at various values of x:

$$d \approx \frac{\Sigma_x \frac{G(x) - I(x)}{I'(x)}}{\Sigma_x 1} \tag{25}$$

By further improvement, the average is weighted using the weighting function defined as:

$$w(x) = \frac{1}{|G'(x) - I'(x)|} \tag{26}$$

The average of the weighting is represented by:

$$d = \frac{\Sigma_x \frac{w(x)[G(x)-I(x)]}{I'(x)}}{\Sigma_x w(x)} \tag{27}$$

The process is repeatedly performed, yielding similar steps to the Newton-Raphson iteration. The estimation ideally converges to the best d. The iteration is expressed by

$$\int_{d_{k+1}=d_k+\frac{\Sigma_x \frac{w(x)[G(x)-I(x+d_k)]}{I'(x+d_k)}}{\Sigma_x w(x)}}^{d_0=0} \tag{28}$$

Feature-based methods are generally more accurate but less robust than intensity-based methods due to its ability to adapt to illumination changes and invariant qualities to scale and rotation. Khorramshahi, Behrad & Kanhere (2008) uses KLT feature detection algorithm to detection OHV yielding favourable results. However, the study did not test in variable weather conditions. KLT can be paired with any of the feature-based methods and tracked using the optical flow methods to determine the direction of OHV to minimise false positive detections. KLT feature tracking assumes that small spatial and temporal changes of motion across an image sequence (Sadeghi-Tehran et al., 2014); therefore, this method works well for static cameras in windy weather conditions.

Previous computer vision approaches to OHVD have proven sensitive to wind. The aims here, to repeat, are to minimise false positive detections caused by camera motion, and to determine the necessary frame and sampling rates required to provide accurate warnings to drivers while comparing various feature detectors. The feature detectors perform similarly, however, they have not been tested for a static camera with minimal movements for OHVD. Therefore, in this chapter, a comparison of the feature detectors is performed.

In the next section, an extension to the proposed approach (mentioned in Chapter 6) is proposed. The extensions include two addition checks: wind analysis and vehicle displacement to minimise the number of false detections. The gap reveals that although these feature detectors have performed well for other vehicle detection scenarios, it has yet to be compared for the OHVD scenario. In this chapter, a comparison of the current state-of-art feature descriptors are explored: 1) SURF (instead of SIFT as it underperforms), 2) BRISK (faster results than SURF and SIFT), 3) EIGEN (Shi and Tomasi's method), 4) HARRIS (classic corner detection algorithm yet robust) and 5) FAST (fast as the name suggests). The best performing feature detector will be used as part of the final validation process in Chapter 8.

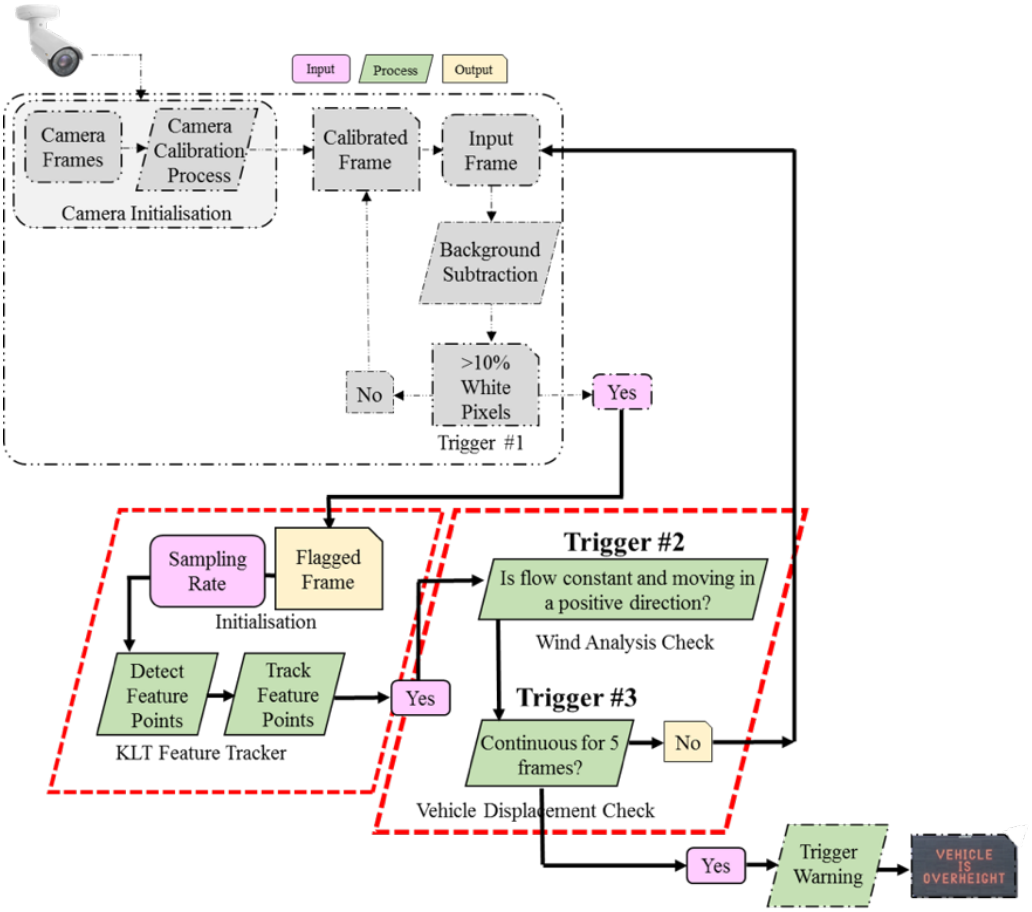


Fig. 73 Extension of proposed approach for over-height vehicle detection using a trigger-based approach.

7.3 Extension of proposed approach

Figure 73 schematises the approach to minimise the number of false positive detections. The OHVD is processed using a trigger-based approach. The first trigger is based on Chapter 6 quantifying the number of white pixels within the ROI. If the number of white pixels exceeds 10%, this passes the frame to Trigger #2. Trigger #2 inputs the sampling rate and uses the KLT feature tracker to determine whether the flow is constant and moving in a positive direction. The need for this check is due to windy weather conditions. If the flow is not constant, the assumption is that the movement is caused by noise, in this case windy weather conditions. To minimise the number of false detections, Trigger #2 is required. The final check (Trigger #3) compares the vehicle displacement in the image. This trigger is analyses the current state-of-art feature detectors to analyse its capabilities to reliably track vehicle displacements over a set number of consecutive frames. If the image passes the three triggers, this will trigger a warning to the driver and flag the image as a positive OH instance. Each step is explained more in detail.

Step 1: The flagged frame is initialised using the input information of the sampling rate. 'Sampling rate' refers to the processing interval between frames, from the point of view of the algorithm. The relationship is to determine the ideal sampling rate in order to sufficiently warn drivers. For example, if a video is recorded at 25 fps, the algorithm can be set to analyse every 1, 5, 10, 30 frames etc., with larger intervals requiring less processing overall. If a vehicle is travelling at 30 mph with a recorded frame rate of 25 fps, the algorithm expects the OH vehicle to be captured in 90 (±5) frames. If sampling is once every 5 frames, the expected positive occurrence is 15–20 times for the same vehicle, equating to same number of positive warnings. This is excessive; only one warning is required for every OH classification, and so the sampling rate needs to be analysed.

Step 2: In this step, optical flow is used to find the physical movement of the 2D projection of the feature points relative to the 2D displacement of pixel patches on the image plane. The initial video frame is a 2D grayscale image, and the feature points are detected using corresponding interest points between a pair of images using local neighbourhoods. The algorithm finds the corners and extracts neighbourhood features. The neighbourhood features are matched and locations of the corresponding feature points are retrieved for each consecutive image. This can be expressed as $(I_i, I_{i+1}, \ldots, I_{i+N})$, where $N <$ number of positive frames.

Step 3: The feature points are initialised and tracked to specify the initial point and video frame location in $[x, y]$ coordinates. The point tracker tracks a set of points using the KLT feature-tracking algorithm from one frame to the next. A block size is initiated to specify the size of the neighbourhood represented by a two-component vector $[height, width]$ around each point being tracked. The neighbourhood corresponds to the spatial matrix area and the block size is set to a minimum to decrease computation time. The output points are an N-by-2 array of $[x, y]$ coordinates that correspond to the new point locations in the subsequent frame. The validity of the points is scored using a confidence scoring technique for each point between 0 (*poor*) and 1 (*perfect*). The scores are a function of the sum of squared differences between the previous and new neighbourhoods. The values correspond to the degree of similarity between the neighbourhood around the previous and new location. Since motion projects to nearby points in the image, spatial coherence and constant feature flow is expected along the x-horizontal dimension of the image.

Step 4: The Wind Analysis Check uses a control variable to analyse the direction of movement of features. This check is based on the characteristics and behaviours of OHV. The tracking of flow and positive directions of movement of objects through frames eliminates the false detections of wind. The Wind Analysis Check is assigned a control variable which analyses the direction of flow and scores each feature-point. Sampling rate information is used as input to activate the KLT feature-tracker detection algorithm. Any motion passing through the ROI is detected and tracked. Each point is tracked over a number of consecutive frames and analysed with reference to its neighbours to determine whether the flow is constant and moving in a positive direction. If the flow is constant, i.e. monotonically

increasing along the x-axis, then a warning is displayed on the OH sign. If the motion is inconsistent and disconnected, the instances are classified as noise, and the process starts over.

The Vehicle Displacement Check evaluates the averages of the positive displacement vectors M to account for camera motion due to wind. For example, let's set the threshold h to a displacement of +5 pixels. This means, if the flow vectors within the ROI is moving in a positive direction at an average of 5 pixels or more, this will increase the counter to $c+1$. If this occurs consecutively for 5 frames or more, this passes the wind analysis #2 and vehicle displacement checks #3 therefore triggering the warning. If the displacement is less than the threshold (meaning less than a displacement of 5 pixels), the counter resets itself to 0.

7.4 Experiment and results

OHV have certain characteristics and behaviours on how they manoeuvre on the roadway and the trajectories that are present. Minimising the number of false positive detections is possible by understanding the behaviours of OHV. The experiment uses the assumption that the camera will endure constant swaying, and that this motion will occur over a significant number of frames (5+) rather than forming an isolated event (1 or 2 frames).

The other assumptions of OHV behaviours are: (1) over a specific height relative to the road plane, (2) travelling in a specific direction (easterly direction at distance d from the camera position) and (3) travelling at a constant speed. The vehicle is assumed to be constant and not decelerating and/or stationary (parked) in front of the camera view. Any motion captured within the ROI is analysed, and can be treated as a binary linear classification problem such that 1=OH vehicle and 0=noise (caused by wind or camera movement):

$$Trigger\ Warning\ System = \begin{cases} 1, overheight\ vehicle \\ 0, noise \end{cases} \tag{29}$$

The limitations of the experiment are evident when vehicles are stopped in front of the camera. This causes a break in the continuity of the vehicle trajectories therefore, invalidating the above assumptions. Minimising these occurrences requires the camera to be installed in such a location away from traffic lights, turning lanes or driveways causing vehicles to slow down.

This section provides details of the data collection, camera specifications and methods used in the experiment, as well as classification accuracy results. The implementation is conducted on 30–50 mph speed limit roadways, which are typical where low bridges exist. The dataset consisted of video involving obstructions such as moving trees or bushes, and building facades with potential background movement, as shown in Figure 74.

Manual swaying of the pole was implemented at times throughout the data collection, in order to simulate wind movement. The dataset consisted of 3661 positive frames, with 102 positive OH occurrences. On average, there were 35.9 frames per every OH occurrence.

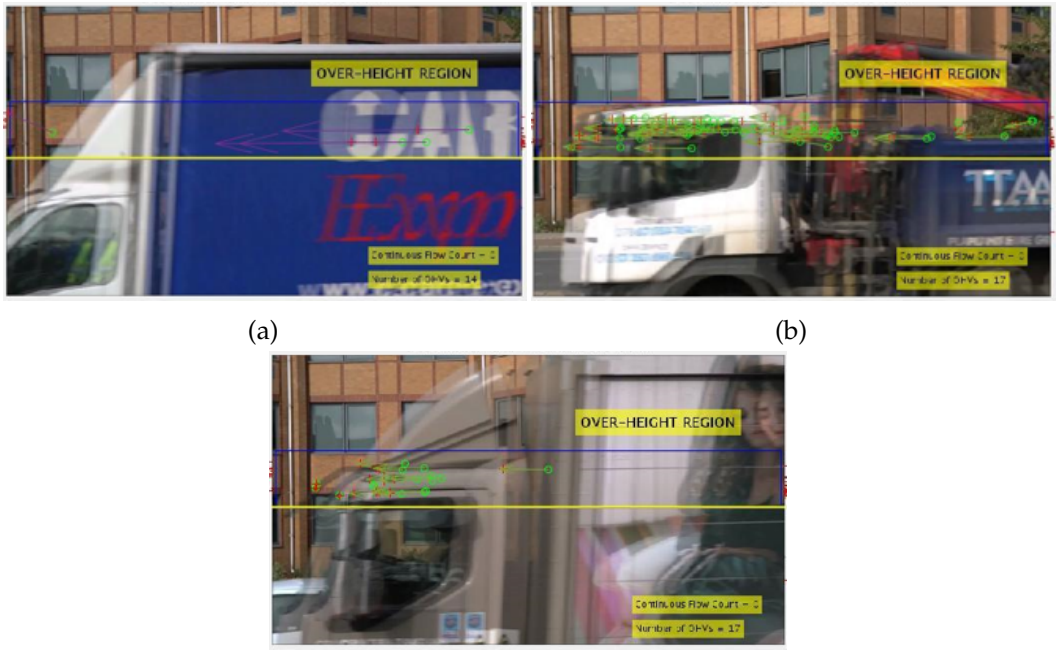
Fig. 74 Examples of the data collected from 6 various locations in the UK and USA.

The experiment uses the same dataset shown in Chapter 6.4. The data is processed using MATLAB R2016a Computer Vision Toolbox on an Intel Core i7-4790. The camera was mounted on a fixed tripod pole where the θ_{yaw} is at 90° and θ_{pitch} at 10°. Although the typical UK low bridge is considered to be 5.03 m (Department of Transport), in this experiment, the camera is set at a height between 2.5 m and 3.2 m to evaluate the flows of the average tall vehicle. The camera was lowered to a height that would allow for a larger dataset. The optimisation parameters revealed in Chapter 6 sets the fixture pixel response threshold (background subtraction) at an intensity level of 142 and the ROI is set at x, y coordinates 70 x 1920. Also, as part of Chapter 6, the trigger was set at 10% (meaning, any white pixels over 10% will trigger an alert to the system to initiate the subsequent steps in the algorithm). OHVD requires a sufficient number of frames to accurately detect the vehicle. The purpose is to distinguish between true OH presence and movement resulting from wind. If the frame and sampling rates are lowered, there is a higher probability of false negatives, whereas at higher rates there are excess frames which increases processing and computing costs. Therefore, an evaluation of the frame and sampling rates is required to determine the optimal parameters.

The experiment was divided into two parts: first to find the ideal sampling and frame rate and second, to find the best feature detector to detect and track vehicle displacements to accurately warn drivers.

Part (1)—Wind analysis check using ideal sampling and frame rate for driver warning

Part (1) includes videos recorded at 25 fps and at 30 fps during windy weather conditions. In order to accurately warn drivers, the ideal sampling and frame rate is needed during windy weather conditions. The experiment is performed to determine the number of positive



(a) (b)

(c)

Fig. 75 Three instances of occlusion when an OHV is traveling in the opposite direction blocking the scene for positive detection.

frames available and the sampling rate required to warn drivers. For each frame rate, the system is evaluated at 5-frame interval sampling rates i.e. every 1, 5, 10, 15, 20, 25 30 frames to determine the optimal sampling rate with sufficient accuracy. Within each sampling rate, the upper and lower limits of the number of positive frames are determined at each frame rate. The number of positive frames is important, as the trigger is set to warn the driver when five consecutive frames have positive motion, i.e. Triggers #2 and #3.

For triggers #2 and #3, these triggers are created based on the behaviours of OHV to minimise the number of false positive detections invoked on the system. For example, trigger #2 calculates the averages of the positive displacement vectors. In Figure 75, there are three instances of vehicles travelling in the opposite direction. The trigger allows the algorithm to disregard objects that are moving in a negative direction and other factors/objects such as birds and wind movement. Although features are detected, the average pixel displacements are not evaluated as the displacement is negative. This allows for better detection and discrimination to minimise the number of false detections.

Trigger #3 is the vehicle displacement check. In Figure 76, five consecutive frames are shown to illustrating the continuous flow counter set on the algorithm. The graph on the bottom left shows the number of tracked feature points over time and the graph on the bottom right shows the number of average displacements over time. In Figure 76c, the graphs are plotted only for this instance of OHV and shows the average detected feature points between 40 and 62 points, while the average positive displacements are between 3

and 6 pixels. Although calculating the average detected feature points is useful for detecting objects present in the scene, the results alone is not sufficiently for accurate detection.

For example, if an average of detected feature points is calculated over time for the dataset, a threshold can be derived. Let's say the threshold is 45 feature points. If the algorithm sets the threshold at 45 feature points, this means that an OHV is present only if the feature points is equal to or greater than 45 feature points. In Figure 76, a plot is shown illustrating the positive OHV occurrences over a period of time (approximately between 5500 and 6000 frames). Each feature point represents a point in time, x, and feature points detected per frame, y.

For Figs. 77a and 77b, it is sufficient to plot a threshold line to indicate an OHV presence vs non OHV presence. However, in Fig 77c, this is no longer the case. When wind is introduced into the scene, the number of detected feature points dramatically increases creating an environment for higher false positive detections. Therefore, triggers #2 and #3 are in place to measure the average positive displacements vectors over 5 consecutive frames. This will eliminate/ minmise the false readings from windy camera movements.

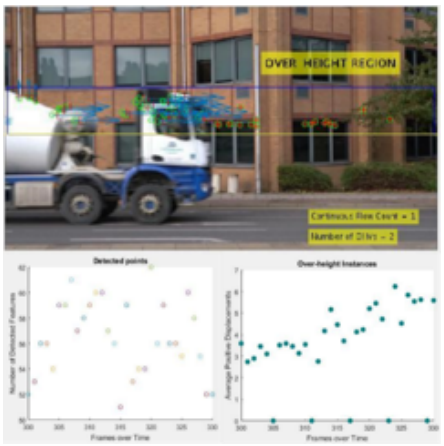
Part (2)—Vehicle displacement check by determining the ideal feature matching method

Optical flow is used to detect and track the motion features evaluating 5 feature detectors: SURF, BRISK, EIGEN, HARRIS and FAST paired with KLT feature-tracker algorithm. A confidence score is assigned to each point to assess validity.

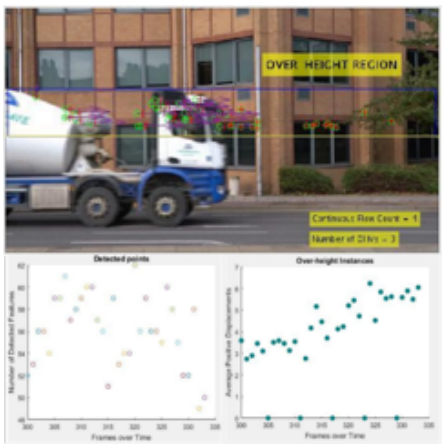
Figure 78 shows the algorithm implemented in MATLAB using the feature matching technique using local neighbourhoods and the Harris algorithm to show the motion of the vehicle. The red vehicle indicates the initial frame indicating the starting position and the cyan vehicle as the sampled frame showing its finishing position; the total motion of the vehicle is shown sampled at every 10 frames.

The feature detector is evaluated for all instances and instances in which it is known to **fail or miss an OHV in the scene**. In Figure 80a, insufficient points cause the algorithm to miss the OHV due to no corners being detectable. In this instance, the vehicle did not pass triggers #2 (wind analysis) and #3 (vehicle displacement) therefore is not classified as an OHV vehicle. Figure 80b shows a tractor vehicle moving through the scene. Out of the 5 feature detectors, only SURF is able to reliably track the vehicle over 5 frames consecutively. The other feature detectors failed due to the average number of positive flow vectors falling below the threshold for detection.

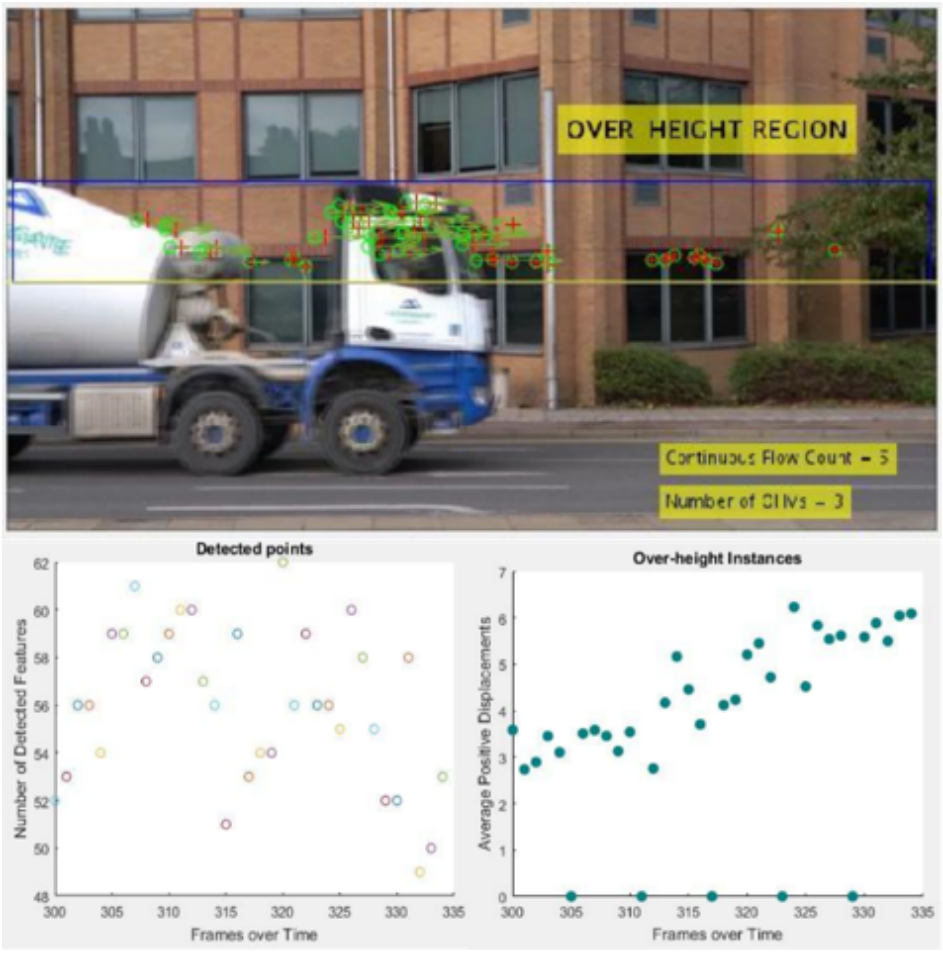
In Figure 79 are shown two instances in which the algorithm did not pick up sufficient features to track adequately for five frames or more. In instance Figure 18a, the vehicle did not pass Trigger #2 where the average displacement of pixels did not pass the threshold limit, resulting in a missed vehicle. In Figure 80b, the bus is detected in the initial and latter



(a) Continuous flow count = 1

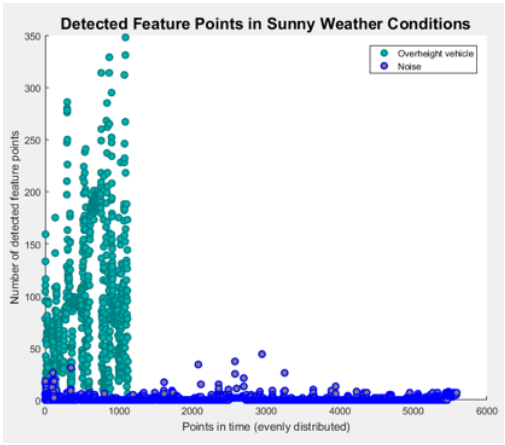


(b) Continuous flow count = 2 . . .

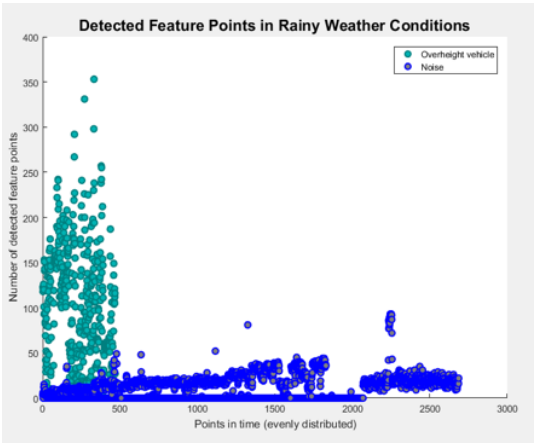


(c) Continuous flow count = 5

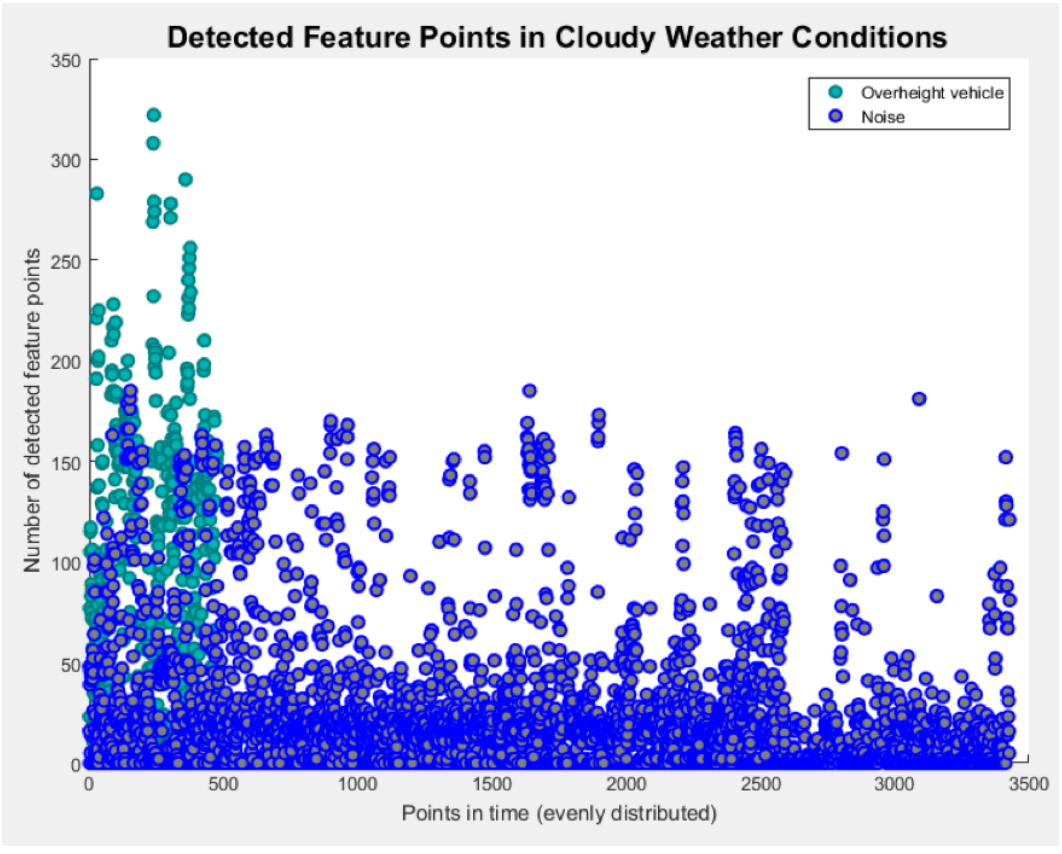
Fig. 76 The five frames show the positive direction of the flow vectors and the increase in the counter when the average pixel displacement is over the set threshold. The graph on the left measures the number of detected feature points while the graph on the right displays the average pixel displacement over time.



(a)



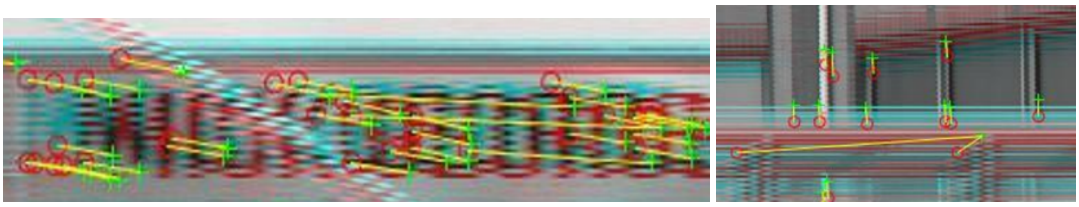
(b)



(c)

Fig. 77 Number of detected features over time in sunny, rainy and cloudy weather conditions.

(a)



(b) (c)

Fig. 78 Over-height vehicle passing through the region of interest with KLT feature vectors detected on the image. The red circles represent the starting pixel location and the green cross represents the ending pixel location. The yellow line represents the length of the pixel movement sampled at every 10 frames (b) zoomed in capture of the KLT vehicle displacements (c) shows the KLT vehicle displacements in windy weather conditions; voiding the assumptions of the algorithm.

(a) (b)

Fig. 79 (a) shows an instance when insufficient features are tracked on the vehicle to pass triggers #2 and #3; (b) shows a tracker vehicle in the scene. Out of the 5 feature detectors, only SURF could reliable track the vehicle over 5 frames consecutively. The other feature detectors failed in this instance.

(a) (b)

Fig. 80 Two instances in which the algorithm did not pick up sufficient features to track adequately for five frames or more. (a) The vehicle did not pass trigger #2 (average displacement of pixels over threshold limit) and therefore is missed as an OHV.

ends of the vehicle. However, when the frame captures only the mid-section of the bus, this causes the feature detector to fail due to no detectable corners in the frame.

Figure 81 shows the number of detected and tracked feature points for each feature detector. A visual comparison can be obtained for the number of displacement vectors calculated for each frame comparison.

Part (1)—Results (Wind analysis check): The sampling and frame rate needed to accurately warn drivers is a 1) sampling rate set at every 10 images and 2) frame rate of 25 frames per second.

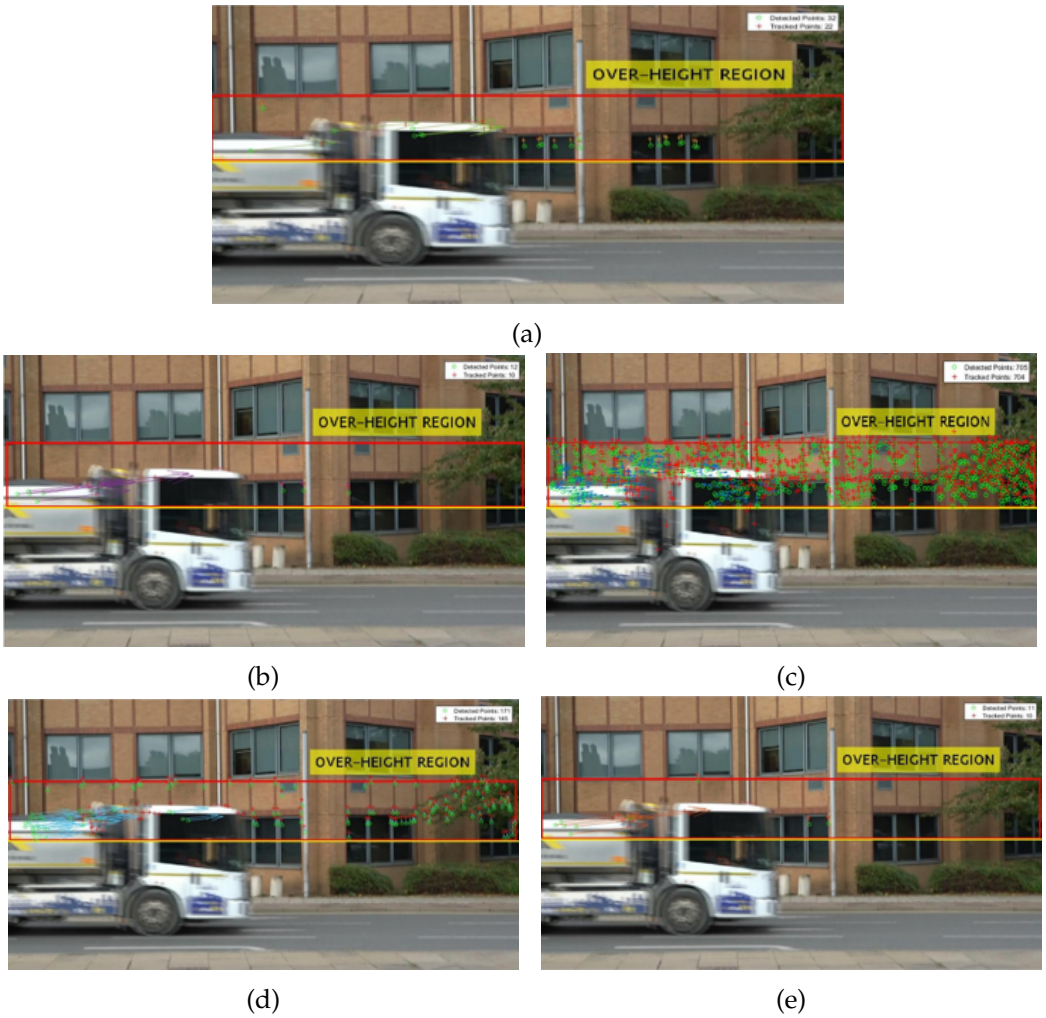


Fig. 81 A frame comparison sampled at 10 frames at 25 fps (frame 0148 and frame 0158 respectively) showing the difference in features detected and tracked for the five feature detectors SURF, BRISK, EIGEN, HARRIS and FAST.

		Number of Positive Frames			Metrics	
	Sampling Rate every x frames	Lower Limit	Average Limit	Upper Limit	Total # of possible warnings per OHV (average)	Actual # of warnings issued to each OHV (average)
25 fps	1 frame	37	49.5	62	9.9	8.5
	5 frames	11	21.5	32	4.3	3.6
	10 frames	7	12	17	2.4	2
	15 frames	3	7.5	12	*	*
30 fps	1 frame	43	57	71	11.4	9.6
	5 frames	17	28	39	5.6	4.6
	10 frames	9	16	23	3.2	2.6
	15 frames	6	7.3	8.5	2.2	0.9
	20 frames	4	6	8	*	*

Table 24 Results recorded at 25 and 30 fps, sampled at x-frame intervals.
**Did not meet minimum required number of positive frames*

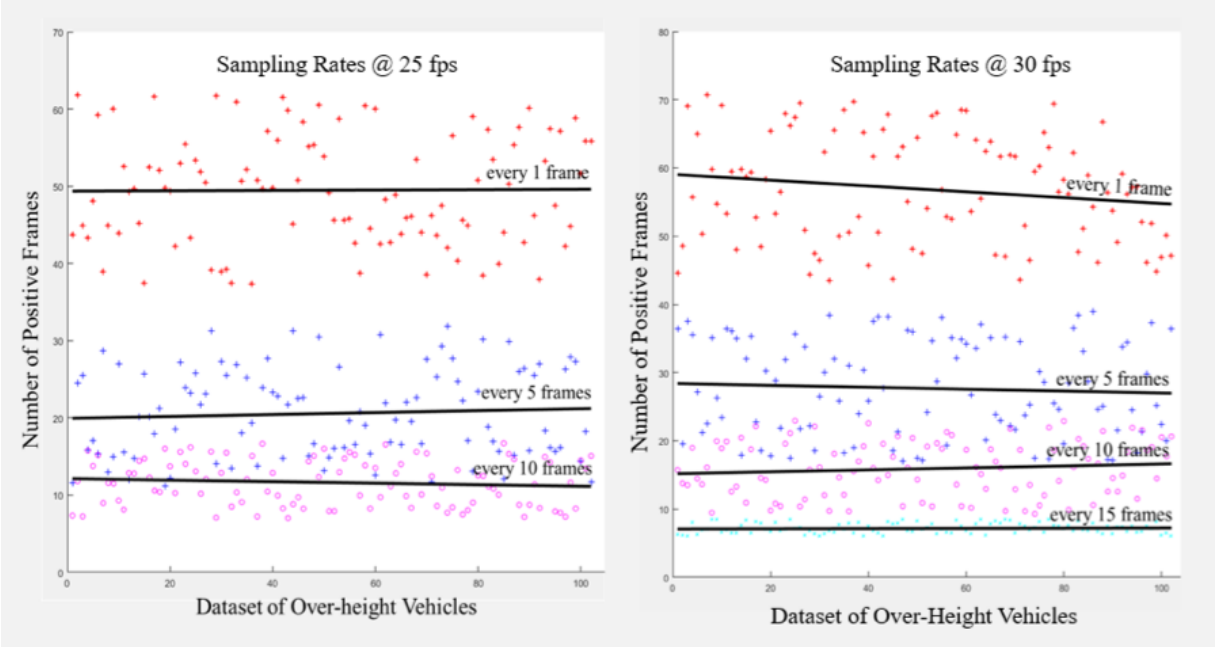


Fig. 82 Average number of positive frames sampled at every 1, 5, 10 and 15 frame intervals recorded at 25 and 30 frames per second.

As shown in Table 24, frame rates are 25 and 30 fps. Figure 82 shows the average of the number of positive frames sampled at every 1, 5, 10 and 15 frame intervals recorded at 25 and 30 frames per second. If the dataset is recorded at 25 fps with a sample rate of every 1 frame, on average, 49.5 positive frames is expected of OH vehicle instances ranging from an upper limit of 62, lower limit of 37 and average of 49.5 positive frames. This results in an average of 8.5 warnings per OH occurrence. Although the algorithm is able to recall positive instances, the multiple warnings decreased the precision to 85.9% compared to an actual # of positive warnings of 9.9. The values replaced by asterisks (*) did not meet the minimum required number of positive frames for consecutive flow analysis and were therefore removed from the dataset. The eliminated values include both 25 and 30 fps video sampled every 15, 20, 25, and 30 frames. The ideal parameters are found to be 25 fps at a sampling rate every 10 frames. The average number of positive frames with these parameters is 12, with an average warning of 2 per OHV.

Part (2) – Results (Vehicle displacement check): SURF feature point detector performs best to track vehicle displacement given frame and sampling rates yielding a recall of 0.941 and precision of 0.879.

In Table 25, the table shows the performance of each of the feature detectors using the 1.9% generality as per Section 6.4. The table shows the results for each of the data collection sites and the false positive (fp), true positive (tp), false negative (fn), and true negative (tn) results.

The total number of frames, total number of OHV and total number of OHV detected are shown.

SITE: UK1	SURF	BRISK	EIGEN	HARRIS	FAST
fp	47	8	603	863	22
tp	394	462	422	436	443
fn	45	13	48	147	25
tn	24231	24237	23633	23245	24202
total frames	24717	24720	24706	24691	24692
total ohv	13	13	13	13	13
ohv detected	14	18	17	25	14
precision	0.893	0.982	0.412	0.336	0.953
recall	0.897	0.973	0.899	0.748	0.947
false positive	0.002	0.000	0.025	0.036	0.001
warning accuracy	0.929	0.722	0.765	0.520	0.929
SITE: UK2	SURF	BRISK	EIGEN	HARRIS	FAST
fp	3	9	162	49	32
tp	306	315	301	169	286
fn	17	41	54	58	89
tn	16759	16717	16589	16825	16691
total frames	17085	17082	17106	17101	17098
total ohv	9	9	9	9	9
ohv detected	9	10	12	13	11
precision	0.990	0.972	0.650	0.775	0.899
recall	0.947	0.885	0.848	0.744	0.763
false positive	0.000	0.001	0.010	0.003	0.002
warning accuracy	1.000	0.900	0.750	0.692	0.818
SITE: UK3	SURF	BRISK	EIGEN	HARRIS	FAST
fp	193	129	144	185	241
tp	386	380	322	203	299
fn	19	25	29	17	77
tn	20292	20350	20423	20511	20289
total frames	20890	20884	20918	20916	20906
total ohv	11	11	11	11	11

ohv detected	12	12	15	18	12
precision	0.667	0.747	0.691	0.523	0.554
recall	0.953	0.938	0.917	0.923	0.795
false positive	0.009	0.006	0.007	0.009	0.012
warning accuracy	0.917	0.917	0.733	0.611	0.917
SITE: UK4	SURF	BRISK	EIGEN	HARRIS	FAST
fp	85	94	132	120	80
tp	881	873	769	691	694
fn	48	63	216	248	325
tn	46515	46497	46579	46484	46378
total frames	47529	47527	47696	47543	47477
total ohv	25	25	25	25	25
ohv detected	26	24	30	37	26
precision	0.912	0.903	0.853	0.852	0.897
recall	0.948	0.933	0.781	0.736	0.681
false positive	0.002	0.002	0.003	0.003	0.002
warning accuracy	0.962	1.042	0.833	0.676	0.962
SITE: UK5	SURF	BRISK	EIGEN	HARRIS	FAST
fp	169	64	84	86	73
tp	1097	1070	976	1005	1064
fn	49	196	479	302	195
tn	59532	59467	59233	59458	59469
total frames	60847	60797	60772	60851	60801
total ohv	32	32	32	32	32
ohv detected	34	32	37	48	33
precision	0.867	0.944	0.921	0.921	0.936
recall	0.957	0.845	0.671	0.769	0.845
false positive	0.003	0.001	0.001	0.001	0.001
warning accuracy	0.941	1.000	0.865	0.667	0.970
SITE: USA1	SURF	BRISK	EIGEN	HARRIS	FAST
fp	23	12	15	2	34
tp	418	397	336	368	366

fn	26	19	31	33	101
tn	22343	22390	22413	22389	22313
total frames	22810	22818	22795	22792	22814
total ohv	12	12	12	12	12
ohv detected	12	11	16	19	14
precision	0.948	0.971	0.957	0.995	0.915
recall	0.941	0.954	0.916	0.918	0.784
false positive	0.001	0.001	0.001	0.000	0.002
warning accuracy	1.000	1.091	0.750	0.632	0.857

Table 25 Performance summary of each feature detector for each scenario:
SURF, BRISK, EIGEN, HARRIS and FAST using a sampling rate at every 10 frames and video recorded at 25 fps.
**Eigen-Shi and Tomasi (1994) method*

SUMMARY	SURF	BRISK	EIGEN	HARRIS	FAST
fp	520	316	1140	1305	482
tp	3482	3497	3126	2872	3152
fn	204	357	857	805	812
tn	189672	189658	188870	188912	189342
total frames	193878	193828	193993	193894	193788
total ohv	102	102	102	102	102
ohv detected	107	107	127	160	110
precision	0.879	0.920	0.747	0.734	0.859
recall	0.941	0.921	0.838	0.806	0.802
false positive	0.003	0.002	0.008	0.009	0.003
warning accuracy	0.958	0.945	0.783	0.633	0.909

Table 26 Performance of each feature detector for all scenarios.
**Eigen-Shi and Tomasi (1994) method*

Table 26 shows the overall performance of each feature detector under all scenarios. The best performing feature detector is SURF, with a recall value of 0.941 and precision value of 0.879. This means that the algorithm is able to detect 94.1% of all OHVS while being able to accurately classify 87.9% of the detected OHV. The total number of OHV is 102. The algorithm detected 107 OHV therefore contributing in the performance of the algorithm. The false positive rate is 0.3%, while the warning accuracy is 95.8% based on the sampling and frame rate parameters. The warning accuracy is calculated based on the total number of OHV divided by the OHV detected by the algorithm. The OHV detected is based on the parameters set by the sampling and frame rate.

7.5 Discussion and conclusions

Part (1) – Windy Analysis Check

The experiment uses a dataset with a total of 102 OH vehicles, recorded at frame rates 25 and 30 fps, respective. At these frame rates, a comparison of sample rates is conducted to determine the sufficient number of positive frames to provide accurate warnings to drivers. The results showed that, for 30 - 50 mph roadways, the optimal frame / sample rate is 25 fps sampled every 10 frames. At this optimal setting, the average number of positive frames is 12. If the trigger is set at five consecutive frames, there is a likelihood of at least one warning to the OH driver, while minimising the number of multiple warnings. As a result, the computing and processing speeds are minimally impaired.

The precision of the algorithm is affected by the number of multiple warnings given for positive OH instances. Therefore, the ideal number to target is 1 warning per OH vehicle with an allowable tolerance of +1. Sampling rates above 15 frames are discarded, as they did not yield the minimum required number of positive frames for processing.

If the trigger is set at a five-consecutive frame minimum, and the lower limit of positive frames is 3, this means that there is a shortage of sufficient frames for processing to meet the five-consecutive frame minimum. Therefore, these sampling rates were not considered further. This configuration is not able to return at least 1 warning due to the low number of frames, which resulted in poor tracking; more frames are required for sufficient processing.

The algorithm is able to minimise the number of misclassifications due to wind by using the behaviours of OH vehicles and windy conditions. However, as a result multiple warnings may be given to a single OH vehicle. An extension of this work is to evaluate and analyse the motion vectors (i.e. vehicle trajectories) to increase the system performance.

Part (2) – Vehicle Displacement Check

A comparison of the feature detectors is conducted to measure its performance using the ideal sampling and frame rate revealed in Part (1). The importance of the feature tracker test is to evaluate the performance of the detector to sufficiently track features of OHV over a consecutive number of frames. In Table 22, SURF performs best with a recall value of 94.1%. Earlier in Chapter 6, the recall target was 95% meaning that only 5% of OHV are to be missed. The best performing feature detector, SURF fell short by 0.9% of the target but despite the margin, is accepted as the ideal feature detector based on the overall results. Ranked second and so forth, BRISK can recall 92.1% of all vehicles while EIGEN, HARRIS and FAST ranking third, fourth and fifth at 83.8%, 80.6% and 80.2% respectively. Overall, the feature detectors performed well yielding a recall disparity of 13.9%, while the precision disparity is less than 2%.

The differences in the detected features play a significant role in the precision and recall results. For example, HARRIS and EIGEN feature detectors detected too many of the corners in the background therefore contributing to the increase in false positive and false negative detections. When wind was present, this rate increased as shown in the results in Table 22. While FAST contributes to a high false negative detection rate due to the misclassifications of the positive frames for negative frames. Instances in which the feature detector failed or missed an OHV are when:

1. A vehicle stops in the scene,
2. Vehicles are moving too slowly or braking in the scene (red light ahead),
3. A vehicle is occluded in the scene (either vehicle travelling in opposite direction or two OHV side by side), and
4. Insufficient features detected (this occurs when let's say, the mid-section of the bus encompasses the entire frame. There are no detectable corners).

These events occur in daily traffic and the results of the feature detectors have considered these occurrences and are reflected in the overall algorithm performance results. To address

cases 1 and 2, there will be events when this occurs. To minimise the risk of non-detection, firstly the camera placement must be carefully thought out as this is critical. The camera placement should be free of upcoming traffic light, stop sign and/or other devices that may stop or queue vehicles. Secondly, if the above is taken into consideration, the alternative is to lower the number of consecutive frames for a positive trigger. If vehicles are moving slower or stopping, the parameter can be set at 2 or 3 for the number of consecutive frames however, as a consequence, this may increase the number of false positive and affect the warning accuracy. Further testing is required.

The last section discusses the false positive and warning accuracy. In Table 3, the performance of the algorithm yields good false positive rates under 0.9% with highest warning accuracy at 95.8% based on the sampling and frame rate parameter settings. The results from Part (1) shows the average number of positive frames is 12 with an actual average warning result of 2 per OHV. At this setting, the warning accuracies performed well for the feature detectors SURF, BRISK, and FAST. With EIGEN and HARRIS, the poor warning accuracies is a result of too many false positive OHV detected, 19.7% more OHV detected for EIGEN and 36.3% more detected for HARRIS. This is due to the sensitivity of the corner detection algorithms resulting in too much noise captured by the feature detectors.

In conclusion, the ideal sampling rate is sampled at every 10 frames at a frame rate of 25 fps while the best performing feature detector yields a recall of 0.941, precision value of 0.879, and warning accuracy of 0.958. Based on the results of Chapters 5, 6 and 7, a final validation test is conducted in Chapter 8 showing the overall performance of the system using the ideal parameters evaluated in the respective chapters.

Chapter 8

Validation of over-height system

In this chapter, the validation of the OHVD and warning system is evaluated as a complete system using the optimised parameters from each of the findings in chapters 5, 6 and 7. The section first reviews the overall framework of the system and devises the parameters into its individual parts.

Secondly, the chapter recaps the systems performance under ideal (sunny, non-windy weather conditions) and the process to minimise the number of misclassifications under non-ideal weather conditions (cloudy, rainy, windy weather conditions). Lastly, the optimised settings are consolidated and validated as a unit. The aim of this section is to:

- **Aim 4:** Validate the overall system performance using the optimised parameters.

In the next section, a recap of the overall framework and validation process is presented.

8.1 Background

In Chapter 5, the system shows proof of concept. In Chapter 6, the system is tested under ideal weather conditions (non-windy) revealing a performance of 99.9% detection with a false positive rate of 0.1%. The parameters included a ROI, threshold value, sampling rate, frame rate and feature detector. In Chapter 7, the same parameters are used and tested using a dataset that consisted more of the typical British weather conditions, sunny, cloudy and rainy weather conditions. The result revealed that under variable weather conditions, the precision decreased by 31.1% to 68.9%, while the algorithm is able to detect 100% of OHV. Initially, the experiment hypothesized that the feature detector could provide a distinctive threshold to classify OHV however, the result revealed that under windy weather conditions, wind plays a large role in the accuracy of the detection and warning to drivers decreasing the accuracy of the system significantly. Additional parameters are added to minimise the number of false positive detections creating further robustness against camera motion.

Prior, each of the parameters are devised into its individual components and tested individually to evaluate and determine its optimal performance setting. In the validation

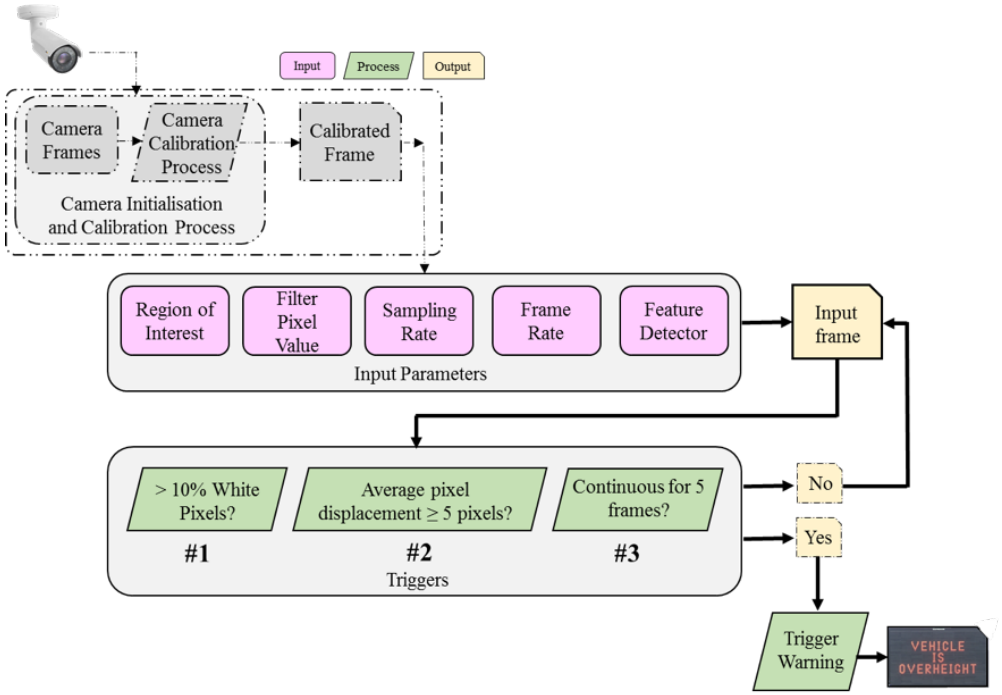


Fig. 83 Validation of over-height vehicle detection and warning system using the five optimised input parameters and three trigger approaches.

Input Parameter				
Region of Interest	Filter Pixel Response Value	Sampling Rate	Frame Rate	Feature Detector
70 vertical x 1920 pixels	142-pixel intensity	Every 10 frames	25 fps	SURF

Table 27 Input parameters used for the validation of the over-height vehicle detection and warning system.

process, the parameters are consolidated and used as inputs into the final experiment. In the next section, the validation approach is described using the optimised parameters.

8.2 Proposed approach

The validation approach shown in Figure 83 is a consolidation of the parameters tested in prior chapters. The five input parameters are outlined in Table 27.

The input parameters used for the final validation step include those parameters optimised in Chapters 6 and 7. Those optimised parameters are used as the five input parameters which makes up the input parameters for validating the OHVD and warning system. The camera system is calibrated and the input frames are passed through a set of triggers. The first trigger is the white pixel threshold, the second trigger is the wind analysis check and

the third trigger is the vehicle displacement check. If the frames pass these three triggers points, a warning is issued to the driver. This is considered a positive OH instance. Each step is explained more in detail below:

Step 1: The first step in the validation approach is to set the parameters in the image using the given input parameters. The first parameter is the ROI with a vertical height of 70 pixels tall by 1920 pixels wide and the filter pixel response value is set to 142-pixel intensity (found in Chapter 6). The sampling and frame rate are set at every 10 frames using the dataset of 25 fps (found in Chapter 7). SRUF feature points is set as the feature detector to detect any corners in the moving objects in the ROI.

Step 2: The input frame passes through three trigger points, otherwise known as checks in the validation process to verify that the object detected behaves as that of a vehicle. Based on the camera calibration, the height of the ROI is calibrated to detect vehicles above 3.4m in this instance.

The first trigger point checks whether the motion within the region exceeds a white pixel intensity difference of 10% or more. If this check is true, 1 is assigned to the image and the image passes onto the next validation trigger. If this is false, the input frame is discarded and the next frame is analysed.

The second trigger is the wind analysis check. The step analyses whether the flow of detected points is constant and moving in a positive direction. This check is based on the characteristics and behaviours of OHV. The tracking of flow and positive directions of movement of objects through frames eliminates the false detections of wind. The Wind Analysis Check is assigned a control variable which analyses the direction of flow and scores each feature-point. Sampling rate information is used as input to activate the KLT feature-tracker detection algorithm. Any motion passing through the ROI is detected and tracked. Each point is tracked over a number of consecutive frames and analysed with reference to its neighbours to determine whether the flow is constant and moving in a positive direction. If the flow is constant, i.e. monotonically increasing along the x-axis, then a warning is displayed on the OH sign. If the motion is inconsistent and disconnected, the instances are classified as noise, and the process starts over.

The Vehicle Displacement Check evaluates the averages of the positive displacement vectors M to account for camera motion due to wind. For example, let's set the threshold h to a displacement of 5+ positive pixels. This means, if the flow vectors within the ROI is moving in a positive direction at an average of 5 pixels or more, this will increase the counter to $c + 1$. If this occurs consecutively for 5 frames or more, this passes the wind analysis (#2) and vehicle displacement checks (#3) therefore triggering the warning. If the displacement is less than the threshold (meaning less than a displacement of 5 pixels), the counter resets itself to 0. If this is false, the input frame is discarded and the next frame is analysed.

Step 3: In this step, the three checks have been triggered and the over-height warning sign is issued to the vehicle driver. The next section carries out the final validation using this process.

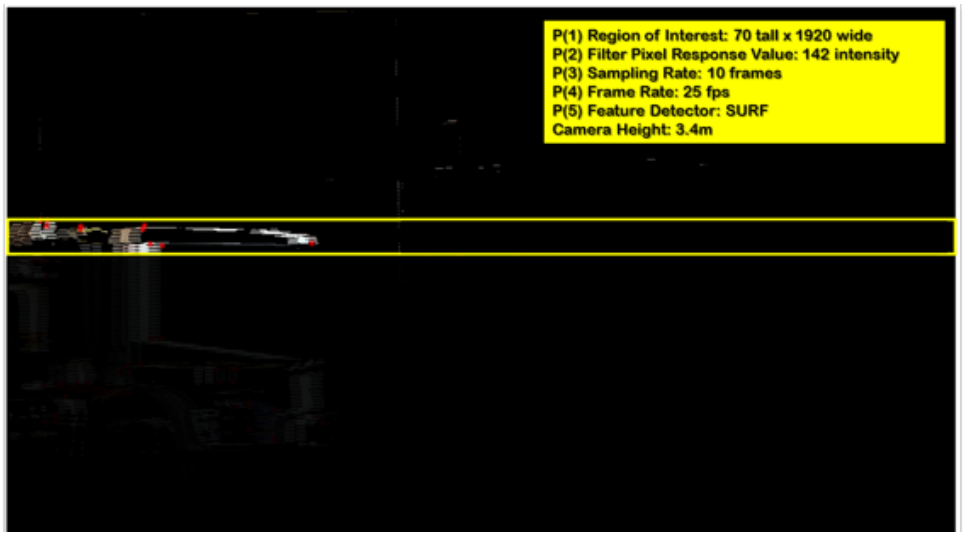


Fig. 84 Validation settings with input parameters 1: Region of Interest, 2: Filter pixel Response Value, 3: Sampling Rate, 4: Frame Rate and 5: SURF Feature Detector used for the final test.

8.3 Experiment and results

This section provides details of the consolidated experiments designed to validate the overall performance of the OHVD and warning system. The implementation is conducted using the same dataset (found in Chapter 6) conducted on two collector roadways with two and four lanes of traffic. The dataset consists of 102 OHV consisting of a generality of 1.9% (3661 positive and 190303 negative frames) recorded at a frame rate of 25 fps in sunny, cloudy and rainy weather conditions. In several scenes, wind is often present.

The aim of the experiment is to validate the overall system performance using the optimised parameters as shown in Figure 84. The optimised parameters are used as inputs into the system, using the three-trigger approach strategy.

Results: Overall system validation performance yielding a recall of 0.945 and precision of 0.911.

The results of the experiment for the six various locations in the UK and USA are presented in Table 28. The validation results show the system able to recall 94.5% of all OHV while accurately classifying each recall frame by 91.1%. The system achieved a low false positive rate at 0.2% while achieving a warning accuracy of 96.6% based on the input sampling and frame rate parameters. In Table 29, the final results for precision and recall are shown showing the modifications in performance as each parameter is optimised throughout the process.

SITE:	UK1	UK2	UK3	UK4	UK5	USA1	Final Validation
fp	27	12	151	7	182	6	385
tp	465	311	387	853	1107	396	3519
fn	32	18	23	61	69	12	215
tn	24190	16767	20330	46615	59392	22378	189672
total frames	24714	17108	20891	47536	60750	22792	193791
total ohv	13	9	11	25	32	12	102
ohv detected	15	10	10	26	33	12	106
precision	0.945	0.963	0.719	0.992	0.859	0.985	0.911
recall	0.936	0.945	0.944	0.933	0.941	0.971	0.945
false positive	0.001	0.001	0.007	0.000	0.003	0.000	0.002
warning accuracy	0.867	0.900	1.100	0.962	0.970	1.000	0.966

Table 28 Results of the validation using the optimised parameters as inputs for the over-height vehicle detection and warning system.

	Description	Precision	Precision Difference	Recall	Recall Difference
Phase 1	Ideal sunny	0.999	—	1.000	—
Phase 2	Variable weather conditions: sunny, cloudy and rainy	0.689	–0.310	1.000	0.000
Phase 3	Minimizing false positive detections by adding three system checks (via trigger points to analyse motion)	0.879	+0.190	0.941	–0.059
Phase 4	Final validation using the optimized parameters from each of the results	0.911	+0.032	0.945	+0.004

Table 29 Final breakdown of the results from each of the phases of testing.

8.4 Discussion and conclusions

The results of the experiment for the six various locations in the UK and USA are presented in Table 28. The validation results show the system able to recall 94.5% of all OHV while accurately classifying each recall frame by 91.1%. The system achieved a low false positive rate at 0.2% while achieving a warning accuracy of 96.6% based on the input sampling and frame rate parameters. In Table 29, the final results for precision and recall are shown showing the modifications in performance as each parameter is optimised throughout the process.

Phase 1: The system is tested under ideal sunny (non-windy) weather conditions resulting in a near perfect performance with a precision of 99.9% and a recall of 100%.

Phase 2: The second phase introduces sunny, cloudy and rainy weather conditions (more British weather you would expect) to assess the performance under these conditions. The system can recall 100% of vehicles, however at a decreased precision of 31% due to the windy weather causing the camera to move. Initially, the number of detected feature points is used as the main trigger to set the alarm, however in windy weather conditions, the feature detector was unreliable and unpredictable, therefore requiring more robust constraints to minimise the false alarm rate. The additional constraints are introduced in Phase 3.

Phase 3: The third Phase adds 'check' points in the form of trigger points to determine whether the motion within the ROI is true to be a vehicle or false due to noise. To implement these triggers, the sampling, frame and feature detector are analysed to determine the best performing parameters. These three parameters play a crucial role in the warning accuracy of the system and the false positive rate. The results of the three parameters decreases the systems recall by 5.9% however, the optimised parameters increased the precision by 19% due to the three checks in place to analyse the motion within the frame.

Phase 4: The final phase validates the system using the five best performing parameters to determine the overall performance, hence the aim of this section. The results reveal a slight drop in recall by 0.4% while the precision accuracy increasing by 3.2%. Overall, the results did not differ largely from Phase 3 and 4 meaning the optimised filter response pixel and the ROI did not play a large factor in its detection. Nevertheless, the initialisation parameters (filter pixel response value and ROI) in Phase 2 still play an important role in the overall detection process for minimised computation speed and cost.

To further enhance the performance of the system while minimising the number of false positive detections, the optical flow and SURF feature detectors performed well with a final recall of 94.5% and precision of 91.1%. To recap, the target recall value is 95-100%. The validation result falls just short of the target by 0.5% however, is minor in the large scheme of the performance. Overall, the OHVD and warning system performs well and can accurately detect potential offending vehicles while maintaining a low false positive rate of

0.2% and warning accuracy of 96.6%. The final chapters conclude with final remarks and contributions to knowledge.

Chapter 9

Conclusions

A bridge or tunnel strike is an incident in which a vehicle, typically a lorry (truck) or double-decker bus, tries to pass under a bridge or tunnel that is lower than its height, subsequently colliding with the structure. In the UK, a low bridge is considered one with a clearance of 5.03m (16′ 6″) or lower. These strikes lead to an increased cost of bridge repairs, clogged up roadways and increased potential for catastrophic events: hazardous spillage and/or total collapse.

The motivation of the research is to develop an affordable yet reliable system to prevent future strikes posing risk to public civil infrastructure. The objective of the research is to create, test and validate a system that is affordable for the average low bridge which outperforms current state of practice. The research hypothesises that implementing a feature-based camera at the height of low clearance bridge will increase the accuracy of OHVD and driver warning. The testing of the system aims to: 1) determine the appropriate calibration process, scene configuration and camera setup, 2) determine the appropriate parameters for image initialisation and detection, 3) further tune the system by minimising the number of false positive detections and lastly 4) to validate the overall performance of the system using the optimised parameters for OHVD. The targets for the system is to achieve a height accuracy of ±5 cm, a recall value of 95–100% and precision of 90%. The experiments are devised into four phases: (1) the camera calibration, (2) image initialisation and optimisation analysis, (3) wind and vehicle displacement analysis and (4) validation.

Chapter 1 describes the problem of bridge and tunnel strikes. The problem may present itself as straight forward at first glance. However, research reveals that the problem is much more complicated and there is no one particular reason why drivers continue to strike into bridges. Network Rail reports this phenomenon occurring every 4.5 hours. The main reasons why these strikes occur are: 1) drivers are unaware of the height of their vehicle, 2) the drivers are equipped with poor route planning and 3) there is inadequate warning of the low structure. The current state of practice that are readily available for consumer purchase falls into three categories: prevention, detection and reporting.

Prevention includes methods to prevent OHVS from occurring and can be subcategorised into passive, sacrificial and active systems. Passive systems are the most common and cost-effective type that exist, due to their readily available supply and 'quick fix' approach; however, passive signing is estimated to be 10-20% effective in preventing incidents. These include signage, variable message signs, bridge markings, and driver educational material such as policies and manuals. Sacrificial systems are physical notification structures such as crash breams (also known as collision protection beams, impact beams, bridge bumpers, or cushion systems), hanging chains, portal frames and road narrowing techniques. The sacrificial systems are estimated to be 30-50% effective, while active systems are EWDS that warn drivers of low approaching bridges. The systems include GPS and optic/laser/light/infrared systems. The active warning systems are estimated to be 50-80% effective and considered most effective at preventing such strikes. Detection methods are used to recognise whether a strike has taken place. These methods include sensors such as accelerometers installed on the bridge structure or other sensing devices to record impact. Reporting are methods to signal a strike to authorities. These methods include close circuit television cameras (CCTV) and wireless remote servers.

The current state of practice in the literature suggests the availability of many effective forms of detection and reporting systems for OHVS that exist, however, the real problem lies in prevention methods. OHV still occurs with high frequency. OHVS active prevention systems (passive, sacrificial and active) on the market are often too expensive that it discourages widespread implementation. Bridge owners are seeking an alternative approach that minimises the occurrences of OHVS. This minimises inspection, maintenance and repair costs.

Chapter 2 discusses the current state of research which shows concentration in prevention, while detection and reporting is already quite efficient. On the prevention side of research, the handful of research falls within computer vision and machine learning techniques. Computer vision and machine learning techniques are a growing area and researchers have found value due to its robustness, computational speed and affordability. Despite their potential significance, vision-based methods have received little attention as a potential solution to the OHVS problem. There has been little research done in computer vision on OHVD, and no real-time implementation of a vision-based system exist.

Chapter 3 discusses a feasibility study conducted in London, United Kingdom in collaboration with Transport for London (London Underground), Cambridge Centre for Smart Infrastructure and Construction (CSIC) and London Borough of Redbridge. The feasibility study uses a modified version of the most recent vision-based method developed at the Georgia Institute of Technology. The method uses a vision-based method to extract the height of box-shaped trucks using a single camera. The results of the feasibility study show that further improvements are needed to cover a larger array of vehicles and overcome the challenges of accurate height detection. The findings of the study remedy the gaps in the

existing literature and report on the results of real-time processing. The lessons learned from the project help developed the proposed solution of this thesis.

Chapter 4 describes the proposed framework that uses a single camera vision-based approach mounted on the side of the roadway. The single camera is able to detect the OHV, capture its number plates and record the scene of the collision in the event of a strike. The same video dataset from the same single camera is used for all three important functions. When the detection occurs, the camera starts to record and the accelerometers on the bridge are activated. If the driver ignores all warning signs and strikes into the bridge, the accelerometers will detect the impact and report this information back to the infrastructure owner(s). The camera system is installed upstream at the height of the low bridge. At that height, the OH plane can be safely approximated as a line in the camera view. Any vehicle exceeding this line is consequently over-height. The camera position and orientation is determined via. a calibration process. Instances of over-height vehicles are detected via. a computer vision detection and tracking algorithm.

The system is designed such that it can easily detect OHV as the visual processing focuses only on offending vehicles therefore improving the computation time and performance of the system. The detection procedure uses motion segmentation as the main feature extraction to detect and track moving objects within a ROI. The detection algorithm calculates the motion differences within the ROI between the current image frame and background model by utilising vehicle motion when OHV are present in the scene. When an OHV is detected, this triggers a warning to the driver.

Chapter 5 describes the experiments designed in Phase 1 to calibrate and test the system to determine the accuracy of the OHV plane in the image. The experiment explores the calibration techniques for setting the OH plane while determining the various camera angles, road speeds and road geometry for optimal camera setup. The basic system is tested under ideal sunny (non-windy) weather conditions resulting in a near perfect performance of 100% recall and 99.9% precision. The false positive rate is 0.1%, and warning accuracy is 100%. The next phase of the testing is meant to test the basic system under more typical British weather (sunny, cloudy and rainy) conditions.

Chapter 6 discusses Phase 2 of the experiments using an optimised procedure to determine the optimal image initialisation parameters for OHVD. The dataset consisted of 102 OHV with a generality of 1.9% (3661 positive, 190303 negative frames) using OHV in the UK and USA on roadways of 25 and 30mph speeds. The optimisation procedure uses three input parameters: (1) filter response value: $cv(threshold)$, (2) verticle window size: $cv(vert)$ and (3) horizontal window size: $dv(horz)$. The two first paramters are variable and the experiment is designed to determine those optimal values. The last variable, dependent on the size of the image is set at 1920 pixels wide. The results of the optimisation converged at a window size of 70 (±3) x 1920 pixels and a filter pixel response value of 142 (±5) pixels. The average precision value is 0.689 (sunny: 0.751; cloudy: 0.631; rainy: 0.685), with recall of 1.000. Although the algorithm can correctly classify 100% of OHV, the *precision* varied

significantly due to wind. As the window size increases, more background noise is detected by the algorithm due to swaying of the pole, leading to a sharp increase in false positive detections. Horizontal sway demonstrated only minimal effects on the 2D OH line projection; however, during windy gusts, more white pixels were captured within the ROI leading to triggering the algorithm. In severe windy conditions, this can be considerably a risk and therefore compromise the system accuracy. Chapter 6 reveals that under ideal weather conditions (low wind, concomitantly, lack of precipitation / other occlusion), the distinction between OHV, wind and other noise is distinguishable by the number of feature points detected. Extrapolating from ideal conditions to *all* weather conditions, the hypothesis is that whenever the number of detected feature points exceeded a threshold value, OHV would indeed have passed through the frame (i.e. the case would be a true positive); however, this proved incorrect. In windy weather conditions, the results proved to be inconclusive resulting in a decrease of performance by nearly 31.0%, all due to false alarms. Phase 3 is designed to minimise these false positive and false negative detections.

Chapter 7 covers Phase 3 of the experiments by adding a further extension designed to assess camera motion and stabilisation under variable weather conditions and validate the overall system using the optimised parameters; an effort to further minimise false positive detection instances. In Phase 3, the experiment is divided into two parts: (1) analyses the wind and (2) vehicle displacement to determine whether the motion vectors can be tracked to differentiate motion (OHV vs. noise) to provide accurate detection and warning to drivers. Triggers are added to the system to act as 'check' points to evaluate the direction and consistency of the motion. If the trajectories do not behave (or move) as an expected vehicle should, the motion is regarded as noise.

Part 1 aims to determine the sampling and frame rate needed to accurately warn drivers and Part 2 compares five feature detections: SURF (speeded-up robust features), BRISK (binary robust invariant scalable keypoints), HARRIS, EIGEN (Shi and Tomasi) and FAST (Features from Accelerated Segment Test) feature detectors to evaluate its performance. The feature detector is paired with optical flow to estimate the motion of the image objects using the KLT Feature Detection Algorithm.

Part 1 of the experiment determined that a sampling and frame rate needed to accurately detect and warn drivers is a sampling rate set at every 10 images and frame rate recorded at 25 frames per second. Part 2 evaluated and analysed the best feature detector to track motion within the ROI and determine the false positive and warning accuracy of the algorithm. The experiment reveals SURF feature detector performs best to track vehicle displacement given frame and sampling rates yielding a recall of 0.941 and precision of 0.879. SURF fell short by 0.9% of the target (95%) but despite the margin, is accepted as the best feature detector based on the overall results. The experiment achieved a low false positive rate of 0.3% and warning accuracy of 95.8%. Chapter 8 and final phase in the thesis, Phase 4 reveals the final validation experiment which consolidates the five optimal parameters and validates it as a unit. The five input parameters are: (1) region of interest, (2) filter pixel

response value, (3) sampling rate, (4) frame rate and (5) feature detector. The optical flow and SURF feature detectors performed well with a final recall of 94.5% and precision of 91.1%. The recall experiences a slight drop of 0.4% while the precision accuracy increasing by 3.2%. Overall, the results did not differ largely from Phase 3 showing the optimised filter response pixel and the ROI did not play a large factor in its detection. The OHVD and warning system performs well and can accurately detect potential offending vehicles while maintaining a low false positive rate of 0.2% and warning accuracy of 96.6%. Nevertheless, the initialisation parameters (filter pixel response value and ROI) in Phase 2 still play an important role in the overall detection process for minimised computation speed and cost. Overall, the OHVD and warning system performs well and can accurately detect potential offending vehicles while maintaining a low false positive rate and warning accuracy. The final section concludes with contributions and future works.

9.1 Contributions and future works

The contribution to research and wider asset infrastructure community is a novel framework to prevent bridge and tunnel strikes using a vision-based solution for over-height vehicle detection for warning drivers. No other vision-based over-height vehicle and detection system exist; this is a first of a kind extensive validation. The main innovation is the single camera setup installed at the side of the roadway at an order of magnitude lower cost than existing systems with minimal compromise of accuracy and performance. The main driving cost for existing systems is the need to install permanent infrastructure.

The proposed system outperforms existing systems on cost due to eliminating the need for new permanent infrastructure. The technology targets the needs of the infrastructure owner and provides a system for asset management through preventative measures to maintain life expectancy of structures and to decrease the cost of repairs, maintenance and inspections. The technology addresses the problem of bridge and tunnel strike prevention and management for improving network capacity (both transport and rail), efficiency (minimising delays and congestion caused by strikes) and operation (for infrastructure owners) through an innovative system designed for urban living and high-volume transport routes. When strikes occur, these accidents represent a great nuisance and threat to public safety leaving the infrastructure owners to remedy the costs. The many benefits of the system are: 1) acts as a means of protecting civil infrastructure from structural damage and minimising disruptions on transport networks, and 2) for warning drivers and preventing injuries, damages and fatalities. The camera setup installation procedure is designed with the infrastructure owner in mind to accommodate multiple laneways of traffic (despite the direction of other traffic), capture the number plates and record the scene of the collision.

The major contributions of the system is the full testing and validation of a vision-based system for OH vehicle detection that bridge and asset owners can use today for bridge and tunnel strike prevention. The system is free to use with no usage restrictions. The camera

is designed with optimised parameters so practitioners can take the camera specifications and hardware requirements and implement at low bridges today. The system is designed to achieve an accuracy of ±5cm within the actual height of the vehicle with a warning accuracy greater than 98% and low false positive detection rate of less than 2%. The system is designed to capture 100% of all OH vehicles with a precision rate of 95%.

The minor contributions of the system are the design and easy set-up and calibration. This can be done by any practitioner using a referencing pole set at the correct height. The dataset is available on GitHub to access the available 102 OH vehicles on request for others to test and analyse. The data was tested on typical traffic within 30-50 mph roadways in variable weather conditions during day- and night time conditions. This is useful for asset owners interested in installing on roadways with similar characteristics. In addition, the single camera is designed to accommodate for multi-lane and multi-directional traffic, which existing prototypes do not have. The system provides a trigger-based approach that can be stream lined to send notifications to asset-owners using web browser access in real-time with minimal maintenance.

Lastly, the camera system can be purchased for less than £3,000, therefore providing asset owners with an affordable, reliable and accurate technology for the prevention against bridge and tunnel strikes at a magnitude less cost than current-state-of-art. As a result, asset owners are able to protect their structures and warn potential offenders while minimising the cost of asset repairs, maintenance and inspections against potential OHVS Recommendations for future works are to expand on the functionalities of the system and to further exploit the computer vision and machine learning techniques. Today, asset owners are moving towards 'smart' intelligent transportation systems. More research can be explored in this area such as the need for statistical data analysis and bridge management classification information. Imagine if the OHVD and warning system could collect useful statistical data and report this information back to the asset owner. This process could save hundreds of man-hours from having to record this information manually and/or retrieving previous reports for information. The system can be equipped with cloud-based features with centralised access for all decision makers to access up-to-date information on the number of detected over-height vehicles, near-misses and prevented bridge strikes. The system can prove its effectiveness and demonstration its true value by feeding statistical data back to the asset owner.

In the event a bridge strike occurs, a bridge management system can be designed to classify the frequency of the vehicle impedance and determine whether the strike is minor, moderate or severe. This information can be fed back to the asset owner for rapid decision maker to: 1) close the bridge, 2) call for emergency, or 3) send an inspector. The bridge management system can be used by the asset owners as guidance to determine the appropriate course of action. The two additions are complimentary and provides the asset owner with more control over their asset infrastructure while providing the capabilities to make decisions using up-to-date readily available information.

Chapter 10

References

Alberta Infrastructure & Transportation. (2008). "Alberta installs over-height warning systems for truckers. Today's Trucking". <http://www.todaystrucking.com/alberta-installs-over-height-warning-systems-for-truckers> (Jun. 17, 2015).

Agrawal, A. K. (2011). *Bridge Vehicle Impact Assessment: Final Report.* University Transportation Research Center, New York State Department of Transportation.

Anagnostopoulos, C. N. E., Anagnostopoulos, I. E., Psoroulas, I. D., Loumos, V., & Kayafas, E. (2008). License plate recognition from still images and video sequences: A survey. *IEEE Transactions on intelligent transportation systems*, 9(3), 377-391.

Baya, H., Essa, A., Tuytelaars, T., & Van Goola, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. *Computer vision–ECCV 2006*, 404-417.

BBC. (2016). "M20 motorway shut after lorry crash causes bridge collapse." BBC News. Kent, United Kingdom. Retrieved from: http://www.bbc.com/news/uk-england-kent-37204050

Betke, M., Haritaoglu, E., & Davis, L. S. (2000). Real-time multiple vehicle detection and tracking from a moving vehicle. *Machine vision and applications*, 12(2), 69-83.

Byrne, A. (2009). *Railway Bridges in Ireland & Bridge Strike Trends.* Railway Safety Commission.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.

Cawley, P. M. (2002). Evaluation of Overheight Vehicle Detection/Warning Systems. In *Today's Transportation Challenge: Meeting Our Customer's Expectations* (No. CD-016).

Chandana, S. (2011). *Real time video surveillance system using motion detection.* In India Conference (INDICON), 2011 Annual IEEE (pp. 1-6). IEEE.

Cheung, S. C. S., & Kamath, C. (2005). Robust background subtraction with foreground validation for urban traffic video. *EURASIP Journal on Advances in Signal Processing*, 2005(14), 726261.

Ching, S. C., & Kamath, C. (2004). Robust techniques for background subtraction in urban traffic video. In *Electronic Imaging 2004* (pp. 881-892). International Society for Optics and Photonics.

Chen, L., El-Tawil, S., Xiao Y., (2016). Reduced Models for Simulating Collisions between Trucks and Bridge Piers. *J. Bridge Eng*, 21(6): 04016020.

Chetverikov, D, Verestóy, J. (1999) Feature point tracking for incomplete trajectories. *Computing*, 62 (4), pp. 321-338.

Cohen, I., & Medioni, G. (1999). Detecting and tracking moving objects for video surveillance. In *Computer Vision and Pattern Recognition, 1999.* IEEE Computer Society Conference on (pp. 319-325). IEEE.

Coifman, B., Beymer, D., McLauchlan, P., & Malik, J. (1998). A real-time computer vision system for vehicle tracking and traffic surveillance. *Transportation Research Part C: Emerging Technologies*, 6(4), 271-288.

Coughlan, J. M., & Yuille, A. L. (2003). The Manhattan World Assumption: Regularities in scene statistics which enable Bayesian inference. *Neural Computation*, 15(5), 1063–1088.

Criminisi, A., Reid, I., & Zisserman, A. (2000). *Single View Metrology*, 40(2), 123–148.

CSS. (2007). *Prevention of Strikes on Bridges over Highways: A Protocol for Highway Managers and Bridge Owners.*

Dai, F., Park, M. W., Sandidge, M., & Brilakis, I. (2015). A vision-based method for on-road truck height measurement in proactive prevention of collision with overpasses and tunnels. *Automation in Construction*, 50, 29-39.

Danker, A. J., & Rosenfeld, A. (1981). Blob detection by relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1), 79-92.

Dawood, M., Cappelle, C., El Najjar, M. E., Khalil, M., & Pomorski, D. (2012). Harris, SIFT and SURF features comparison for vehicle localization based on virtual 3D model and camera. In *Image Processing Theory, Tools and Applications (IPTA)* (pp. 307-312). IEEE.

Department of Transportation UK. (2008). *Traffic Signs Manual - Regulatory Signs* (Chapter 3).

Department of Transportation UK. (2007). "Control of vehicle (rules 117 to 126), Rule 126 Stopping Distances." Retrieved from: https://www.gov.uk/guidance/the-highway-code/general-rules-techniques-and-advice-for-all-drivers-and-riders-103-to-158

Department of Transportation UK. (2013). *Traffic Signs Manual - Warning Signs* (Chapter 4).

Department of Transportation London, UK. (1990). "Bridge Strike Statistics." Retrieved from: https://www.gov.uk/government/organisations/department-for-transport

Do, B., (2009), MATLAB, Image Blur Metric, source code: http://uk.mathworks.com/matlabcentral/fileexchange/24676-image-blur-metric/content/blurMetric.m

Donaldson, J. H. (1906). *U.S. Patent No. 810,041.* Washington, DC: U.S. Patent and Trademark Office.

Donnelley, R. (2010). "Strike it out: Preventing Bridge Strikes (Vol. 35)." Retrieved from: https://www.transport.gov.scot/media/26732/j12054.pdf

Dutch Ministry of Infrastructure, & Environmental Department of Waterways and Public Works. (2015). "Netherlands deploys 'pictorial' truck-height warning system at Velsertunnel, Traffic Technology Today". <http://www.traffictechnologytoday.com/news.php?NewsID=66219> (Jun. 17, 2015).

Elgammal, A., Duraiswami, R., Harwood, D., & Davis, L. S. (2002). Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of the IEEE*, 90(7), 1151-1163.

Elgammal, A., Harwood, D., & Davis, L. (2000). Non-parametric model for background subtraction. *Computer Vision* (pp.751-767). ECCV.

El-Tawil, S., Severino, E., & Fonseca, P. (2005). Vehicle Collision with Bridge Piers. *Journal of Bridge Engineering ASCE*, 3(345).

Fathi, H., & Brilakis, I. (2014). Multistep explicit stereo camera calibration approach to improve euclidean accuracy of large-scale 3D reconstruction. *Journal of Computing in Civil Engineering*, 04014120.

Federal Highway Administration (FHWA). (2013). “Interstate System Conditions and Performance”. <http://www.fhwa.dot.gov/infrastructure/intrstat.cfm> (Jun. 17, 2015).

Fontaine, M. D. (2003). Engineering and Technology Measures to Improve Large Truck Safety: State of the Practice in Virginia. Retrieved from: http://ntl.bts.gov/lib/37000/37100/37130/03-tar13.pdf

Friberg, J. O. (2013). Detecting background and foreground from video in real-time with a moving camera. *IMINT Image Intelligence.* Retrieved from: http://uu.diva-portal.org/smash/get/diva2:640050/FULLTEXT01.pdf

Fu, C. C., Burhouse, J. R., & Chang, G. L. (2004). Overheight vehicle collisions with highway bridges. *Transportation Research Record: Journal of the Transportation Research Board*, 1865(1), 80-88.

Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2009, June). Manhattan-world stereo. In *Computer Vision and Pattern Recognition, 2009.* CVPR 2009. IEEE Conference on (pp. 1422-1429). IEEE.

Galer, M. (1980). An ergonomics approach to the problem of high vehicles striking low bridges. *Applied ergonomics*, 11(1), 43-46.

Ghose, A. (2009). Strategies for the management of bridges for vehicular impacts. *Proceedings of the ICE-Structures and Buildings*, 162(1), 3-10.

Gorur, P., & Amrutur, B. (2011). Speeded up gaussian mixture model algorithm for background subtraction. In: *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on* (pp. 386-391). IEEE.

Gupte, S., Masoud, O., Martin, R. F., & Papanikolopoulos, N. P. (2002). Detection and classification of vehicles. *Intelligent Transportation Systems, IEEE Transactions on*, 3(1), 37-47.

Haering, N., Venetianer, P. L., & Lipton, A. (2008). The evolution of video surveillance: an overview. *Machine Vision and Applications*, 19(5), 279-290.

Hanchey, C. M., & Exley, S. F. (1990). Overheight Vehicle Warning Systems in Missippi. *ITE journal*, 60(6).

Hardik, J. H., Shah, N. N., & Raval, D. P. (2013). Automated Video Surveillance Sytem for Human Motion Detection. *Journal of Information, Knowledge and Research in Electronics and Communication Engineering*, 2(2), 675–678.

Harris, C., & Stephens, M. (1988). A combined corner and edge detector. In *Alvey vision conference* (pp. 10-5244).

Heikkila, J., & Silvén, O. (1997). A four-step camera calibration procedure with implicit image correction. In *Computer Vision and Pattern Recognition, 1997.* (pp. 1106-1112). IEEE.

Hinz, S. (2005). Fast and subpixel precise blob detection and attribution. In *Image Processing, 2005, IEEE International Conference on* (pp. III-457). IEEE.

Hocenski, Z., Vasilic, S., & Hocenski, V. (2006). Improved canny edge detector in ceramic tiles defect detection. In *IEEE Industrial Electronics, IECON 2006-32nd Annual Conference on* (pp. 3328-3331). IEEE.

Horberry, T., Halliday, M., & Gale, A. G. (2002). Bridge strike reduction: optimising the design of markings. *Accident Analysis & Prevention*, 34(5), 581-588.

Huang, K. Y., Tsai, Y. M., Tsai, C. C., & Chen, L. G., (2010). "Video stabilization for vehicular applications using SURF-like descriptor and KD-tree", *2010 17th IEEE International Conference on Image Processing (ICIP 2010)*, pp. 3517-3520.

Huijsmans, D. P., & Sebe, N. (2005). How to complete performance graphs in content-based image retrieval: Add generality and normalize scope. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(2), 245-251.

Ismail, K., Sayed, T., & Saunier, N. (2013). A methodology for precise camera calibration for data collection applications in urban traffic scenes. *Canadian Journal of Civil Engineering*, 40(1), 57-67.

Jazayeri, A., Cai, H., Zheng, J. Y., & Tuceryan, M. (2011). Vehicle detection and tracking in car video based on motion model. *IEEE Transactions on Intelligent Transportation Systems*, 12(2), 583-595.

Jiang, Y., Xu, Y., & Liu, Y. (2013). Performance evaluation of feature detection and matching in stereo visual odometry. *Neurocomputing*, 120, 380-390.

Jog, A., & Halbe, S. (2013). Video Analytics Trip Wires for Surveillance System. *International Journal of Computer Applications*, 65(8).

Kanhere, N. K., & Birchfield, S. T. (2010). A taxonomy and analysis of camera calibration methods for traffic monitoring applications. *IEEE Transactions on Intelligent Transportation Systems*, 11(2), 441-452.

Khorramshahi, V., Behrad, A., & Kanhere, N. K. (2008). Over-Height Vehicle Detection in Low Headroom Roads Using Digital Video Processing. *World Academy of Science, Engineering and Technology*, 287–291.

Kim, K., Chalidabhongse, T. H., Harwood, D., & Davis, L. (2005). Real-time foreground–background segmentation using codebook model. *Real-time imaging*, 11(3), 172-185.

Laservision (2014). Softtop Laservision Australia. Retrieved from: http://www.laservision.com.au/page.asp?lid=1&sec=Projects&subsec=Permanent+Attractions&subsubsec=Softstop

Lavanya, M. P. (2014). Real time motion detection using background subtraction method and frame difference. *Int. J. Sci. Res.(IJSR)*, 3(6), 1857-1861.

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011). BRISK: Binary robust invariant scalable keypoints. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 2548-2555). IEEE.

Liu, W., Chen, S., & Hasuer, E., (2012). Bridge Clearance Evaluation Based on Terrestrial LIDAR Scan. *J. Perform. Construction Facility.* 26:469-477.

London Underground, 2013, personal communication, Ashok Parmar, Senior Planner (Civil Maintenance), October 2013.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.

Lowry, E. G., & Forster, L. N. (1981). U.S. Patent No. 4,284,971. Washington, DC: U.S. Patent and Trademark Office.

Lucas, B. D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. (pp. 674-679).

Manjunathan, K., Albert, S., & Deeter, D. (2005). A Summary of Rural Intelligent Transportation Systems (ITS) Benefits as applied to ODOT Region 1, (April).

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, 207(1167), (pp. 187-217).

Martin, A., & Mitchell, J. (2004). *Measures to Reduce the Frequency of Over-Height Vehicles Striking Bridges: Final Report* (Unpublished). Transport Research Laboratory (Vol. T/079/04). Crowthorne: Berkshire.

Martínez-Martín, E., & Del Pobil, A. P. (2012). *Robust motion detection in real-life scenarios.* Springer Science & Business Media.

Massoud, M. A. (2013). Over-height Vehicle Detection System in Egypt. *World Congress on Engineering, II*, (pp. 3–6).

Mattingly, S. P. (2003. Mitigating Overheight Vehicle Crashes into Infrastructure: A State of the Practice. In *Proceedings of the 82nd Annual Meeting of the Transportation Research Board, Washington, DC.*

Mehrani, E., Ayoub, A., & Ayoub, A. (2009). Evaluation of fiber optic sensors for remote health monitoring of bridge structures. *Materials and Structures*, 42(2), 183-199.

Meyer, G. (2013). Texas Department of Motor Vehicles. CVISN plans for bridge-hit warnings. Retrieved from: http://txdmv.gov/publications-carriers/doc_download/3371-mcd-dispatch-fall-winter-2013

Mirzaei, F. M., & Roumeliotis, S. I. (2011). Optimal estimation of vanishing points in a manhattan world. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 2454-2461). IEEE.

Mittal, A., & Paragios, N. (2004). Motion-based background subtraction using adaptive kernel density estimation. In *Computer Vision and Pattern Recognition, 2004. Proceedings of the 2004 IEEE Computer Society Conference on* (Vol. 2, pp. II-302). IEEE.

National Weather Service (2017). UK wind map. Retrieved from: https://www.metoffice.gov.uk/public/weather/wind-map

Network Rail. (2007a). *Prevention of Strikes on Bridges over Highways: A Protocol for Highway Managers and Bridge Owners* (p. 40).

Network Rail. (2007b). *Prevention of Bridge Strikes: A Good Practice Guide for Transport Managers.* London.

Network Rail (2017). The risk of bridge strikes. Retrieved from: https://www.networkrail.co.uk/running-the-railway/looking-after-the-railway/bridges-tunnels-viaducts/risk-bridge-strikes/

New York State. (2015). "Governor Cuomo Announces Installment of Over-Height Vehicle Detection Systems to Improve Roadway Safety on the Hutchinson River Parkway". <https://www.governor.ny.gov/news/governor-cuomo-announces-installment-over-height-vehicle-detection-systems-improve-roadway> (Jun. 17, 2015).

Nguyen, B. & Brilakis, I. (2016). Understanding the Problem of Bridge and Tunnel Strikes Caused by Over-height Vehicles. *Transportation Research Procedia*, 14. 3915-3924.

Nguyen, B. & Brilakis, I. (2017). Over-Height Vehicle Detection: Minimising misclassifications due to wind. In *2nd Lean Construction (LC3) conference. Session: Intelligent Transportation Systems (JC3).* 4th – 12th of July 2017. Heraklion, Crete, Greece.

Nguyen, B., Brilakis, I., & Vela, P. A. (2016). Vision-Based Over-Height Vehicle Detection. In *Transportation Research Board 95th Annual Meeting (No. 16-3550)*. 8th of January 2016. Washington, DC, USA.

Nguyen, B., Brilakis, I., & Vela, P. A. (2017). Optimized parameters for over-height vehicle detection under variable weather conditions. In *Journal of Computing for Civil Engineering*, 31(5), 04017023.

Niu, L., & Jiang, N. (2008). A moving objects detection algorithm based on improved background subtraction. In *Intelligent Systems Design and Applications, 2008. ISDA'08. Eighth International Conference on* (pp. 604-607). IEEE.

Olufs, S., & Vincze, M. (2011a). Robust single view room structure segmentation in manhattan-like environments from stereo vision. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on* (pp. 5315-5322). IEEE.

Olufs, S., & Vincze, M. (2011b). Towards robust Room Structure Segmentation in Manhattan-like Environments from dense 2.5 D data. In *Control, Automation and Systems (ICCAS), 2011 11th International Conference on* (pp. 1491-1496). IEEE.

Park, M.-W., Dai, F., Sandidge, M. J., & Brilakis, I. (2013). Vision-Based Approach for Measuring On-Road Truck Heights. In *Creative Construction Conference*.

Pathan, I., Chauhan, C., Kathiriya, P., Scholar, P. G., & Professor, A. (2016). Real Time Moving Object Detection and Tracking Using Adaptive Gaussian Mixer Model and Lucas-Kanade Method. *International Journal of Engineering Science*, 4888.

Qiao, P., Yang, M., & Mosallam, A. S. (2004). Impact analysis of I-Lam sandwich system for over-height collision protection of highway bridges. *Engineering Structures*, 26(7), 1003-1012.

Rosten, E., & Drummond, T. (2005). Fusing points and lines for high performance tracking. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* (pp. 1508-1515). IEEE.

Rosten, E., Porter, R., & Drummond, T. (2010). Faster and better: A machine learning approach to corner detection. *IEEE transactions on pattern analysis and machine intelligence*, 32(1), 105-119.

Sadeghi-Tehran, P., Clarke, C., & Angelov, P. (2014). A real-time approach for autonomous detection and tracking of moving objects from UAV. In *Evolving and Autonomous Learning Systems (EALS), 2014 IEEE Symposium on* (pp. 43-49). IEEE.

Sandidge, M. J. (2012). Truck height determination using digital video (MSc dissertation, Georgia Institute of Technology).

Scotland, T. (2011). bridgestrikes. Retrieved from: https://bridgestrikes.wordpress.com/

Services, Roads & Maritime. (2013). Overheight vehicles - Is your vehicle overheight?. Retrieved from rms.nsw.gov.au

Shah, M. (2014). Impaired driving charge in Burlington Skyway crash. Toronto Sun. Toronto. Retrieved from: http://www.torontosun.com/2014/07/31/burlington-skyway-lanes-closed-due-to-crash

Shanafelt, G. O., & Horn, W. B. (1980). Damage evaluation and repair methods for prestressed concrete bridge members. *NASA STI/Recon Technical Report N81*, 24313.

Shanafelt, G. O., & Horn, W. B. (1984). *Guidelines for evaluation and repair of damaged steel bridge members (No. HS-037 759)*.

Shao, J., Zhou, S. K., & Chellappa, R. (2010). Robust height estimation of moving objects from uncalibrated videos. *IEEE Transactions on Image Processing*, 19(8), 2221-2232.

Sharma, H., Hurlebaus, S., & Gardoni, P. (2008). Development of a bridge bumper to protect bridge girders from overheight vehicle impacts. *ComputerAided Civil and Infrastructure Engineering*, 23(6), 415-426.

Sharma, H., Hurlebaus, S., & Gardoni, P. (2012). Performance-based response evaluation of reinforced concrete columns subject to vehicle impact. *International Journal of Impact Engineering*, 43, 52-62.

Shi, J., & Tomasi, C. (1994). Good Features to Track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on* (pp. 593-600). IEEE.

Shneier, M. (1983). Using pyramids to define local thresholds for blob detection. *IEEE transactions on pattern analysis and machine intelligence*, (3), 345-349.

Sina. (2012). "Over-height warning device safeguards tunnel." <http://english.sina.com/china/2012/1207/535183.html> (Jun. 17, 2015).

Solutions, F. S. (2017). Overheight Detection and Barrier. Retrieved from: http://futurenetsecurity.com/products/access-control/overheight-detection/

Song, G., Olmi, C., & Gu, H. (2007). An overheight vehicle–bridge collision monitoring system using piezoelectric transducers. *Smart materials and structures*, 16(2), 462.

Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on* (pp. 246-252). IEEE.

Stauffer, C., & Grimson, W. E. L. (2000). Learning patterns of activity using real-time tracking. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8), 747-757.

Strainstall. (2017). BridgeWatch®. Retrieved from: https://www.strainstall.com/files/2314/6607/9893/Strainstall_Bridge_Monitoring_Brochure_V1.0.pdf

Sturm, P. F., & Maybank, S. J. (1999). On plane-based camera calibration: A general algorithm, singularities, applications. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. (pp. 432-437). IEEE.

Tomasi, C., & Kanade, T. (1992). Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2), 137-154.

Thakuriah, P., & Geers, D. G. (2013). *Transportation and information: trends in technology and policy*. New York: Springer.

Trivedi, M. M., Gandhi, T. L., & Huang, K. S. (2005). Distributed interactive video arrays for event capture and enhanced situational awareness. *IEEE Intelligent Systems*, 20(5), 58-66.

Tsai, R. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation*, 3(4), 323-344.

University of California at Santa Cruz, D. of C. E. (2017). Image Features: Edges and Corners, CMPE 264: Image Analysis and Computer Vision. Retrieved December 5, 2017, from https://classes.soe.ucsc.edu/cmpe264/Fall06/Lec5.pdf

Unzueta, L., Nieto, M., Cortés, A., Barandiaran, J., Otaegui, O., & Sánchez, P. (2012). Adaptive multicue background subtraction for robust vehicle counting and classification. *IEEE Transactions on Intelligent Transportation Systems*, 13(2), 527-540.

Urazghildiiev, I. R., Ragnarsson, R., Wallin, K., Rydberg, A., Ridderstrom, P., & Ojefors, E. (2002). A vehicle classification system based on microwave radar measurement of height profiles. *RADAR, IEEE* (pp. 409–413).

Urazghildiiev, I., Ragnarsson, R., Ridderstrom, P., Rydberg, A., Ojefors, E., Wallin, K., Enochsson, P, Ericson M & Lofqvist, G. (2007). Vehicle classification based on the radar measurement of height profiles. *IEEE Transactions on intelligent transportation systems*, 8(2), 245-253.

Valgren, C., & Lilienthal, A. J. (2010). SIFT, SURF & seasons: Appearance-based long-term localization in outdoor environments. *Robotics and Autonomous Systems*, 58(2), 149-156.

Vargas, M., Milla, J. M., Toral, S. L., & Barrero, F. (2010). An enhanced background estimation algorithm for vehicle detection in urban traffic scenes. *IEEE Transactions on Vehicular Technology*, 59(8), 3694-3709.

von Gioi, R. G., Jakubowicz, J., Morel, J. M., & Randall, G. (2010). LSD: A fast line segment detector with a false detection control. *IEEE transactions on pattern analysis and machine intelligence*, 32(4), 722-732.

Washington State Department of Transportation. (2013). "I-5 Skagit River Bridge Replacement". <http://www.wsdot.wa.gov/projects/i5/skagitriverbridgereplacement/> (Jun. 17, 2015).

Wikipedia, 2017. (2017). Pyramid (image processing). Retrieved from: https://en.wikipedia.org/wiki/Pyramid_(image_processing)

Wojek, C., Roth, S., Schindler, K., & Schiele, B. (2010). Monocular 3d scene modeling and inference: Understanding multi-object traffic scenes. In *European Conference on Computer Vision* (pp. 467-481). Springer Berlin Heidelberg.

Xu, L. J., Lu, X. Z., Smith, S. T., & He, S. T. (2012). Scaled model test for collision between over-height truck and bridge superstructure. *International Journal of Impact Engineering*, 49, 31-42.

Xu, N., Rangwala, S., Chintalapudi, K. K., Ganesan, D., Broad, A., Govindan, R., & Estrin, D. (2004, November). A wireless sensor network for structural monitoring. In *Proceedings of the 2nd international conference on Embedded networked sensor systems* (pp. 13-24). ACM.

Yang, M., & Qiao, P. (2010). Analysis of cushion systems for impact protection design of bridges against overheight vehicle collision. *International Journal of Impact Engineering*, 37(12), 1220-1228.

Yang, Y., Chen, J. X., & Beheshti, M. (2005). Nonlinear perspective projections and magic lenses: 3D view deformation. *IEEE Computer Graphics and Applications*, 25(1), 76-84.

Yoo, Y., & Park, T. S. (2008, June). A moving object detection algorithm for smart cameras. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on* (pp. 1-8). IEEE.

Zhang, G, Avery, R. P., Wang, Y, (2007) Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras. *Transp. Res. Rec., vol. 1993*, pp. 138-147.

Zheng, Q., & Chellappa, R. (1995). Automatic feature point extraction and tracking in image sequences for arbitrary camera motion. *International journal of computer vision*, 15(1-2), 31-76.

Zivkovic, Z., & Van Der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7), 773-780.